Renormalized Quantum Yang-Mills Fields in Curved Spacetime

Stefan Hollands*

School of Mathematics
Cardiff University
UK
and
Institut für Theoretische Physik
Universität Göttingen
FRG

Dedicated to K. Fredenhagen on the occasion of his 60th birthday

February 28, 2008

*HollandsS@Cardiff.ac.uk

Abstract

We present a proof that quantum Yang-Mills theory can be consistently defined as a renormalized, perturbative quantum field theory on an arbitrary globally hyperbolic curved, Lorentzian spacetime. To this end, we construct the non-commutative algebra of observables, in the sense of formal power series, as well as a space of corresponding quantum states. The algebra contains all gauge invariant, renormalized, interacting quantum field operators (polynomials in the field strength and its derivatives), and all their relations such as commutation relations or operator product expansion. It can be viewed as a deformation quantization of the Poisson algebra of classical Yang-Mills theory equipped with the Peierls bracket. The algebra is constructed as the cohomology of an auxiliary algebra describing a gauge fixed theory with ghosts and anti-fields. A key technical difficulty is to establish a suitable hierarchy of Ward identities at the renormalized level that ensure conservation of the interacting BRST-current, and that the interacting BRST-charge is nilpotent. The algebra of physical interacting field observables is obtained as the cohomology of this charge. As a consequence of our constructions, we can prove that the operator product expansion closes on the space of gauge invariant operators. Similarly, the renormalization group flow is proved not to leave the space of gauge invariant operators. The key technical tool behind these arguments is a new universal Ward identity that is formulated at the algebraic level, and that is proven to be consistent with a local and covariant renormalization prescription. We also develop a new technique to accomplish this renormalization process, and in particular give a new expression for some of the renormalization constants in terms of cycles.

Contents

1 Introduction **5**

1.1 Generalities 5
1.2 Renormalization of theories without local gauge invariance 7
1.3 The problem of local gauge invariance 8
1.4 Summary of the report 10
1.5 Guide to the literature 11

2 Generalities concerning classical field theory **12**

2.1 Lagrange formalism 12
2.2 Yang-Mills theories, consistency conditions, cohomology 17
2.3 Proof of the Algebraic Poincare Lemma, and the Thomas Replacement Theorem 26

3 Quantized field theories on curved spacetime: Renormalization **32**

3.1 Definition of the free field algebra $\mathbf{W}_0$ for scalar field theory 32
3.2 Renormalized Wick products and their time-ordered products 36
3.3 Inductive construction of time-ordered products 41
3.4 Examples 55
3.5 Ghost fields and vector fields 60
3.6 Renormalization ambiguities of the time-ordered products 62
3.7 Perturbative construction of interacting quantum fields 65

4 Quantum Yang-Mills theory **69**

4.1 General outline of construction 69
4.1.1 Free fields 69
4.1.2 Interacting fields 71
4.1.3 Operator product expansions and RG-flow 73
4.2 Free gauge theory 76
4.3 Interacting gauge theory 80
4.4 Inductive proof of Ward identities T12a, T12b, and T12c 84
4.4.1 Proof of T12a 96
4.4.2 Proof of T12b 100
4.4.3 Proof of T12c 104
4.5 Formal BRST-invariance of the *S*-matrix 104
4.6 Proof that $d\mathbf{J}_I = 0$ 105
4.7 Proof that $Q_I^2 = 0$ 107
4.8 Proof that $[Q_I, \Psi_I] = 0$ when Ψ is gauge invariant 111
4.9 Relation to other perturbative formulations of gauge invariance 112

5 Summary and outlook **115**
5.1 Matter fields, anomalies 117
5.2 Other gauge fixing conditions 117
5.3 Background independence 118

A U(1)-gauge theory without vector potential **119**

B Effective Actions in curved spacetime **123**

C Wave front set and scaling degree **125**

D Hadamard parametrices **127**
D.1 Scalar Hadamard parametrix 127
D.2 Vector Hadamard parametrix 131

E Hadamard states **131**

1 Introduction

The known interactions of elementary particles seem to be well-described by quantized field theories with local gauge invariance such as QCD. Such theories have been extensively investigated in the context of flat Minkowski spacetime from a variety of different angles. It has in particular been demonstrated that these quantum field theories are internally consistent, at least to all orders in the renormalized perturbation expansion. The early Universe on the other hand is described by a strongly curved spacetime, and important new quantum field theory effects arise in this situation— an important example being the generation of primordial fluctuations that have left an imprint in the CMB as well as the large scale structure of the universe. For this reason, it is obviously important to study quantum gauge theories in curved Lorentzian spacetimes such as the expanding Universe. The question how to consistently construct such theories in arbitrary curved, globally hyperbolic spacetimes is an open problem.

As a first step in this direction, we will prove in this paper that perturbative non-abelian pure Yang-Mills theory can be consistently quantized on any globally hyperbolic spacetime, to all orders in perturbation theory, and any gauge group G that is a direct product of $U(1)^l$ and a semi-simple Lie group. The essence of our proof is the inductive construction of an explicit renormalization prescription for the perturbatively defined interacting field quantities that preserves gauge invariance, and that depends locally and covariantly upon the spacetime metric. The proof of this statement is rather complicated, and it relies partly on auxiliary constructions that have been previously given in the literature. Some of these constructions are not so widely known as the renormalization techniques in flat spacetime, and there is at present no comprehensive review. We therefore found it appropriate to present these constructions in the form of a report.

1.1 Generalities

Quantum field theory in curved spacetime is a natural generalization of flat space quantum field theory in which one considers quantized fields propagating on a rigidly fixed, non-dynamical, Lorentzian spacetime rather than flat Minkowski spacetime. In order to have a well-defined propagation of such fields (even at the classical level), one usually assumes that the spacetime does not have any gross causal pathologies such as closed time-like curves, (a typical assumption is that the spacetime is "globally hyperbolic") but otherwise no restrictions on the metric are placed. In particular, one does not have to (and does not want to) assume that the metric has any isometries, or that it is a solution to a particular field equation. As quantum field theory on flat spacetime, quantum field theory on curved spacetime is in general only believed to be an effective theory with a limited range of validity. It is expected to loose predictive power when the spacetime curvatures become as large as the inverse Planck length, or in quantum states where typical quantum field observables such as the quantum stress energy operator have expectation values or variances (fluctuations) of the order of the Planck length. On the other hand, the theory is expected to be a very good approximation when the spacetime curvatures are of

the order (or below) the scale of elementary particle physics such as Λ_{QCD}, or even the grand unification (GUT) scale, which is expected to be the relevent scale during inflation. Naturally, it is also in this regime (as well as in the case of black holes) that the most interesting physical effects predicted by the theory occur.

Independent of those questions regarding the limits of physical applicability of quantum field theory in curved spacetime, one may ask whether this theory, in itself, has a consistent mathematical formulation or not—just as it is a relevant question whether classical mechanics has a well-defined mathematical formulation even though it clearly has a limited range of validity as a physical theory. Unfortunately, this question is a very difficult one, which has not been answered in a satisfactory manner for interacting quantum field theory models even in flat spacetime (in 4 dimensions). Nevertheless, there exist perturbative approaches to interacting quantum field theory in Minkowski spacetime, and it is by now well-understood how to calculate, in principle, terms of arbitrary high order in the perturbation expansion. In particular, one has a good understanding how to systematically deal with the problem of renormalization that needs to be addressed at each order to get meaningful expressions, and it is known how to calculate quantities of physical interest for, say, the purposes of collider physics. In fact, this approach is at present by far the most powerful method to obtain theoretical predictions for particle physics experiments, and to test quantum field theory.

In quantum field theories in curved spacetime, new conceptual problems arise because one no longer has a preferred vacuum state in time-dependent spacetimes, as may be understood from the familiar fact that time-dependent background fields tend to give rise to particle creation. Thus, a state that may be thought of as a vacuum at one time may fail to be the vacuum at later time. This suggests to use an S-matrix formulation of the theory, but such a formulation also does not make sense in general if the spacetime does not have any asymptotically time-independent regions in the far past or future, or if the metric approaches a time-independent metric too slowly. At the technical level, one no longer has a clear cut relation between quantum field theory on Lorentzian spacetimes and Riemannian spacetimes, because a general (even analytic) Lorentzian spacetime will not be a real section in a complexified manifold that also has a real, Riemannian section. Furthermore, familiar flat space techniques such as momentum space, dimensional regularization, the Euclidean path integral, are not available on a curved manifold.

As had been realized for some time, these conceptual problems can in principle be overcome by shifting the emphasis to the local quantum field operators, which can be unambiguously defined on any (globally hyperbolic) Lorentzian spacetime. The key insight was that the algebraic relations between the quantum fields (such as commutators, or the "operator product expansion") have an invariant meaning for any such spacetime, even if there are no states with a definite particle interpretation. Nevertheless, it remained an unsolved problem how to construct in practice interesting (non-free) quantum field theories perturbatively on a general globally hyperbolic spacetime, mainly because of the very complicated issues related to renormalization on a curved manifold. A fully satisfactory construction of perturbative, renormalized quantum field theory on curved space was finally given in a series of papers [18, 17, 62, 63, 64] where it was

shown that the algebras of local observables (interacting local fields) can always be constructed at the level of formal power series in the coupling, independent of the asymptotic behavior of the metric at infinity. It was shown in detail how to perform the renormalization process in a local and covariant way, and it was thereby seen that the remaining finite renormalization ambiguities correspond to the possibility of adding finite local terms (possibly with curvature couplings) to the Lagrangian, and to the possibility of making finite field-redefinitions ("operator mixing with curvature"). These constructions also provided a completely new, geometrical understanding of the nature of the singularities of multi-point operator products and their expectation values in terms of "microlocal analysis" [72, 16, 95], and thereby provided a geometric generalization of the usual spectrum condition in Minkowski spacetime quantum field theory to curved manifolds. By considering the behaviour of the theory under a rescaling of the metric $g \to \mu^2 g$, a definition of the renormalization group could be given [64], and detailed results about the (poly-logarithmic) scaling behavior of products of interacting field operators were thereby obtained. It is also understood how to construct the operator product expansion from the algebra of interacting fields in curved space, and this gives direct information about the interplay between quantum field interactions and spacetime curvature at small scales [68].

1.2 Renormalization of theories without local gauge invariance

The building blocks in the renormalized perturbation series for the interacting fields are the time-ordered products $T_n(\mathcal{O}_1 \otimes \cdots \otimes \mathcal{O}_n)$ of composite fields in the underlying free field theory. In standard approaches in flat spacetime, these objects are typically viewed as operators on a Hilbert space ("Fock-space"), but in curved spacetime there is no preferred Hilbert-space representation. In this context, it is more useful to view them instead as members of an abstract algebra, which may in the end be represented on a Hilbert space (typically in infinitely many inequivalent ways). The first step in the renormalization program therefore is to define a suitable abstract algebra, and this can indeed be done using the techniques of the "wave front set." The next step is to actually construct the time-ordered products as specific elements in this algebra. A naive definition leads to infinite meaningless expressions, but one can show that it is possible to obtain meaningful objects by a process called "renormalization". Conceptually, the best approach here is to first formulate a set of conditions ("renormalization conditions") on the time-ordered products to be constructed, and then show via an explicit construction that these properties can be satisfied. It turns out that the conditions do not uniquely fix the time ordered products, but there remain certain finite renormalization ambiguities. In curved spacetime, it is a major challenge to formulate sufficiently strong renormalization conditions in order to guarantee that these ambiguities only consist in adding finite "contact terms" at each order n, which are covariant expressions of the Riemann curvature and the fields of a suitable dimension. A key condition to guarantees this is that the T_n should themselves be local and covariant [62], and a precise formulation of that condition naturally leads to a formulation of quantum field theory in the language of category theory [19]. The condition of locality and covariance is a rather strong one, and it is correspondingly non-trivial to find a renormalization method that will ensure that

this condition is indeed satisfied. Such a scheme was found in [62, 63] for interacting scalar field theory, based on key earlier work of [18, 17], and also on the work [37, 38], where an algebraic variant of perturbation theory in flat space was developed. We will present these constructions in section 3 of the paper. Here we follow the general steps proposed in these references, but we develop a new technique to perform the actual renormalization (extension) step. Our new method (described in the proof of Lemma 6) is more explicit than previous constructions, and also gives an interesting new formula for some of the renormalization constants describing the departure from homogeneous scaling in terms of an integral of a closed form of a cycle in $\mathbb{R}^{4n}$, see Proposition 1.

In quantum field theory, one typically wants certain fields to have special properties. For example, an important observable in any theory with a metric is the stress energy tensor, which is conserved at the classical level if the metric is the only background field (as we assume). One would like the corresponding quantum field to be conserved as well. In perturbative quantum field theory, it is far from obvious that the corresponding interacting quantum field quantity is also conserved, and indeed there exist theories where this fails to be the case [2]. In general, one can formulate a set of renormalization conditions on the time-ordered products (the "principle of perturbative agreement" [66]) that will guarantee conservation to all orders in the perturbation expansion. In [66], it was shown that the question whether or not these identities can be satisfied is equivalent to the question whether a certain cohomological class on the space of all metric defined by the field theory is trivial or not. The obstruction sometimes cannot be lifted, and then the renormalization condition is impossible to satisfy: There are anomalies. Similarly, in gauge theories, one wants certain currents to be conserved at the quantum level and it is important to ensure that there are no anomalies.

1.3 The problem of local gauge invariance

In fact, the perturbative construction of renormalized field theories on curved space without local gauge invariance does not carry over straightforwardly to theories with local gauge invariance, and the construction of such models was therefore up to now an important open problem. The key obstacle is that the field equations of local gauge theories, such as e.g. the pure Yang-Mills theory studied in this paper, are not globally hyperbolic in nature even if the underlying spacetime is globally hyperbolic. This, however, is a basic assumption in the constructions [17, 18, 62, 63]. In theories with local gauge invariance, the field equations fail to be hyperbolic in nature precisely due to local gauge invariance, because it implies that solutions to the field equations are not entirely determined by their initial data on some Cauchy surface as required by hyperbolicity, but also on an arbitrary choice of local gauge. At the classical level, this problem can be dealt with by simply fixing a suitable gauge. However, at the quantum level, it is problematical to base the theory on a gauge-fixed formulation, because gauge fixing typically has non-local features. This causes severe problems e.g. for the renormalization process. An elegant and very successful approach circumventing these problems is the BRST-method [9, 10]. This method consists in replacing the original action by a new action containing

additional dynamical fields. That new action yields hyperbolic field equations, and has an invariance under a nilpotent so-called "BRST transformation", s, on field space. Gauge invariant field observables are precisely those that are annihilated by s, or more precisely, the cohomology classes of s. Furthermore, the classical Poisson (or Peierls) brackets [91, 87, 27, 37] of the gauge fixed theory are invariant under s. Thus, as first suggested by [39] (based on [84]), one can try to proceed by first quantizing the brackets of the gauge fixed action (in the sense of deformation quantization [37, 38, 7, 8]), promote the differential s to a graded derivation at the quantum level leaving the quantized brackets invariant, and then at the end define the algebra of physical observables to be the kernel (or rather cohomolgy) of the quantum BRST-differential. As we will prove in this paper, this program can be carried out successfully for renormalized Yang-Mills theory in curved spacetime, at the level of formal power series in the coupling constant.

Thus, the first step consists in finding an appropriate gauge fixed and BRST invariant modified action, S, for pure Yang-Mills theory in curved space involving the gauge field, and new auxiliary fields ("anti-fields"). This step is completely analogous to Yang-Mills theory in flat space. Next, one needs to "quantize" the brackets associated with the new action S. It is not known presently how to do this non-perturbatively even in flat space, but one can proceed in a perturbative fashion as in theories without local gauge invariance.

The final step special to gauge theories is now to define a quantum BRST derivation acting on the quantum interacting fields This derivation should (a) leaves the product invariant, (b) square to 0, and which (c) go over to the classical BRST transformation s in the classical limit. The natural strategy for constructing the quantum BRST transformation is to consider the quantum Noether current corresponding to the classical BRST-transformation. One then defines a corresponding charge, and defines BRST-derivation via the graded commutator in the star-product with this charge. While this definition automatically satisfies (a), it is highly non-obvious that it would also satisfy properties (b) and (c). In fact, it is even unclear whether that the quantum Noether current operator associated with the BRST-transformations would be conserved, as would be required in order to yield a conserved charge.

The basic reason why it is a non-trivial challenge to establish conservation of the quantum BRST current, as well as (b) and (c), is that the construction of the time ordered products T_n used to define the interacting quantum fields via the Bogoliubov formula involve renormalization. It is far from obvious that a renormalization prescription exists such that interacting BRST current would be conserved, and such that (b) and (c) would hold. In fact, as we will show, these properties follow from a new infinite hierarchy of Ward identities for the time-ordered products [see eq. (336) for a generating functional of these identities], which are violated for a generic renormalization prescription. We will show that there nevertheless exists a renormalization prescription compatible with locality and covariance such that these Ward identities are satisfied in curved space, to all orders in the renormalized perturbation expansion, when the gauge group is a product of $U(1)^l$ and a semi-simple group. Thus, we can define an algebra of interacting quantum fields as the cohomology of the quantum BRST-differential, and this defines perturbative quantum Yang-Mills theory. In a second step, we then define quantum states

(i.e., representations) of this algebra by a deformation argument. Here we rely on a construction invented in [39]. As a by-product of our constructions, we can also show that the operator product expansion in curved space [68] closes among gauge-invariant operators, and that the renormalization group flow likewise closes among gauge-invariant operators.

Our approach has several virtues also in the context in flat spacetime. The key virtue is that, since our constructions are entirely local, there is a clear separation between issues related to the ultra-violet (UV) and infra-red (IR) behavior of the theory. In particular, in our approach, the identities reflecting gauge invariance may be formulated and proved entirely independently from the infrared behavior of the theory, while the infra-red cutoff is only removed in the very end in an entirely well-defined manner at the algebraic level ("algebraic adiabatic limit" [18]). In this way, infra-red divergences are neither encountered at the level of the interacting field algebras, nor in fact at the level of quantum states, i.e., representations[1]. In this respect, our approach is different from traditional treatments based on Feynman diagrams or effective actions, which are only formal in as far as the treatment of the IR-problems are concerned. We explain in some more detail the relation of our approach to those treatments in sec. 4.9.

A local approach that is similar to ours in spirit has previously been taken in the context of QED on flat spacetime in [39], and in [36, 35] for non-abelian gauge theories on flat spacetime. Note, however, that the "Master Ward identity" expressing the conditions for local gauge invariance in [35] was taken as an axiom and has not been shown to be consistent yet[2], as opposed to the Ward identities of our paper, which are shown to hold. Also, our Ward identities (336) appear to be different from those expressed in the Master Ward Identity of [35, 36].

1.4 Summary of the report

This report is organized as follows. In section 2, we first review basic notions from classical field theory, including classical BRST-invariance and associated cohomological constructions. The material in this secion is well-known and serves mainly to set up the notations and provide basic results that are needed in later sections. In section 3, we review the perturbative construction of interacting quantum field theory on curved spacetime. We focus on theories without local gauge invariance. We explicitly describe scalar field theory, and we briefly mention the changes that have to be made for ghost and vector fields (in the Lorentz gauge). We give a detailed renormalization prescription for the time-ordered products, their renormalization ambiguities, and describe how interacting fields may be constructed from them. We also show how the method works in some concrete examples. The material presented in this section is to some extent taken from [18, 62, 63, 37, 36, 34], but there are also some important new developments. In section 4, we perturbatively construct renormalized quantum Yang-Mills theory. We first give an outline of the basic strategy, and then fill in the technical details in the later sections.

[1]However, we would encounter the familiar infra-red divergences if we were to try to construct scattering states. Actually, it is clear that those types of states cannot be defined in a generic curved spacetime anyway even for massive fields, so we do not see this as a problem.

[2]For recent progress in analyzing the validity of the Master Ward identity, see [15].

We present our new Ward-identities in subsection 4.3, and then prove them in section 4.4. We prove in 4.5 that our identities formally imply the BRST-invariance of the S-matrix, in 4.5 that they imply the conservation of the interacting BRST-current, and in 4.6 that they imply the nilpotency of the interacting BRST-charge operator. We conclude and name open problems in section 6. Appendix A contains a treatment of free $U(1)$-theory avoiding the introduction of the vector potential and an explanation of the new superselection sectors arising in this context. The appendices B–E contain definitions and various constructions that are omitted from the main part of the paper.

1.5 Guide to the literature

A standard introduction to the theory of quantum fields on a curved space is [108], which gives an in-depth discussion of the conceptual problems of the theory, as well as the Hawking and Unruh-effect, at the level of free quantum fields. The generalization of the latter effect to certain black-hole spacetimes—emphasizing especially the role of the so-called "Hadamard condition"—is discussed in the review-style article [80]. Other monographs are [49, 11]. The perturbative construction of interacting scalar quantum field theories on curved spaces was given in the series of papers [18, 17, 62, 63, 66]. Important contributions to the understanding of Hadamard states in terms of microlocal analysis, which were a key input in these papers, were made by Radzikowski [93, 94]. These results are reviewed and extended in the very readable paper [77]. A complete characterization of the state space of perturbative quantum field theory using microlocal analysis is given in [65]. A definition and analysis of the renormalization group in curved space was given in [64]. The generalization of the Wilson operator product expansion in curved spacetime was constructed to all orders in perturbation theory in [68]. Perturbative scalar quantum field theory on Riemannian spaces was treated in [20] using the BPHZ method, and by [82] using the method of flow equations. General theorems about quantum field theory in curved spacetime within a model-independent setting were obtained in [67] (PCT-theorem), and by [107] (spin and statistics theorem). The literature on the quantization of gauge theory, and especially Yang-Mills theory in flat spacetime is huge. The use of ghost fields was proposed first by [45], and the early approaches to prove gauge invariance at the renormalized level used the method of Feynman graphs, together with special regularization techniques [69, 70, 71]. More recent discussions based on the Hopf-algebra structure behind renormalization [21, 22, 83] may be found in [105, 106]. With the discovery of the BRST-method [9, 10], cohomological methods were developed and used to argue that gauge invariance can be maintained at the perturbative level in flat spacetime. Comprehensive reviews containing many references are [25, 90, 60, 5], see also e.g. [103, 104, 41, 42, 43]. There are also other approaches to quantum gauge invariance in flat space, based on the Epstein-Glaser method [44] for renormalization. These are described in the monographs [96, 97] and also in [102], which also contain many references. For a related approach, see [101]. The idea to formulate quantum gauge theory at the level of observables, and to implement the gauge invariance in the operator setting was developed in flat space in [39, 36, 35], building on earlier work

of [84]. A somewhat more detailed comparison between the various approaches to the gauge invariance problem and our solution is given in Sec. 4.9, where additional references are given.

2 Generalities concerning classical field theory

2.1 Lagrange formalism

Most, though not all, known quantum field theories have a classical counterpart that is described in terms of a classical Lagrangian field theory. This is especially true for the gauge theories studied in this paper, so we collect some basic notions and results from Lagrangian field theory in this subsection that we will need later. Not surprisingly, for perturbative quantum field theories derived from a classical Lagrangian, many formal aspects can be formulated using the language of classical field theory, but we emphasize that, from the physical viewpoint, quantum fields are really fundamentally different from classical fields.

To specify a classical field theory on an n-dimensional manifold M, we first need to specify its field content. We will generally divide the fields into background fields, collectively denoted Ψ, and dynamical fields, collectively denoted Φ. Both background and dynamical fields are viewed as sections in a certain fibre bundle, $B \to M$, over the spacetime manifold. We will assume that the background fields always comprise a Lorentzian metric $g = g_{\mu\nu} dx^\mu dx^\nu$ over M (which is a section in the bundle of non-degenerate symmetric tensors in $T^*M \otimes T^*M$ of signature $(-++\cdots+)$). More generally, the background fields may comprise a non-abelian background gauge connection, or varions external sources. We will also admit Grassmann-valued fields, which are described in more detail below. The dynamical fields will typically satisfy equations of motion, which are derived from an action principle. By contrast, the background fields will never be subject to any equations of motion.

To set up an action principle, we need to specify a Lagrangian. The Lagrangians that we will consider have the property that they are locally and covariantly constructed out of the dynamical fields Φ, and the background fields Ψ. In particular, they do not depend implicitly on additional background structure such as the specification of a coordinate system. Since such functionals will play an important role in perturbation theory, it is worth defining the notion that a quantity is locally and covariantly out of a set of dynamical and non-dynamical fields Φ, Ψ with some care. Let us denote by $B \to M$ the "total bundle" in which the dynamical and non-dynamical fields live. For example, in case all the fields are tensor fields, the total bundle is simply the direct sum of all the tensor bundles corresponding to the various types of fields. If $x \in M$, we let $J^k_x(B)$ denote the space of "k-jets" over M. This is defined as the equivalence class of all sections $\sigma = (\Phi, \Psi) : M \to B$, with the equivalence relation $\sigma_1 \sim \sigma_2$ if $\nabla^q \sigma_1|_x = \nabla^q \sigma_2|_x$ for all $q \le k$, where ∇ is any affine connection in the bundle B, and where we have put

$$\nabla^k \sigma = dx^{\mu_1} \otimes \cdots \otimes dx^{\mu_k} \nabla_{(\mu_1} \cdots \nabla_{\mu_k)} \sigma . \tag{1}$$

We say that a p-form $\mathcal{O} = \mathcal{O}_{\mu_1 \dots \mu_p} dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_k}$ is constructed out of $\sigma = (\Phi, \Psi)$ and its first

k derivatives if $\mathcal{O}$ is a map

$$\mathcal{O} : J^k_x(B) \to \bigwedge^p T^*_x M \tag{2}$$

for each $x \in M$, which we will also write as $\mathcal{O}(x) = \mathcal{O}[\sigma(x), \nabla\sigma(x), \dots, \nabla^k\sigma(x)]$. Now let $\psi : M \to M'$ be an immersion that lifts to a bundle map $B \to B'$ denoted by the same symbol, and let σ and σ' be sections in $B \to M$ respectively $B' \to M'$ such that $\sigma = \psi^*\sigma'$. We will say that $\mathcal{O}$ is a p-form that is locally constructed out of the fields σ if we have

$$\mathcal{O}[\sigma(x), \nabla\sigma(x), \dots, \nabla^k\sigma(x)] = \psi^*\mathcal{O}[\sigma'(x'), \nabla\sigma'(x'), \dots, \nabla^k\sigma'(x')], \quad \psi(x) = x', \tag{3}$$

for any x and any such embedding ψ. This condition makes precise the idea that $\mathcal{O}$ is only constructed out of $\sigma = (\Phi, \Psi)$ and finitely many of its derivatives, but depends on "nothing else". For example, if the fields are a background metric, g, and a set of dynamical tensor or spinor fields Φ, then one can show that $\mathcal{O}$ can depend upon the metric only via the curvature, i.e., it may be written in the form

$$\mathcal{O}(x) = \mathcal{O}[\Phi(x), \nabla\Phi(x), \dots, \nabla^k\Phi(x), g(x), R(x), \nabla R(x), \dots, \nabla^{k-2}R(x)] \tag{4}$$

where ∇ is now the Levi-Civita (or spin-) connection associated with g, and $R = R_{\mu\nu\sigma\rho}(dx^\mu \wedge dx^\nu) \otimes (dx^\sigma \wedge dx^\rho)$ is the curvature tensor. This result is sometimes called the "Thomas replacement theorem," and a proof may be found in [76]. The second example relevant to this work is when the background fields contain in addition a background gauge connection ∇ in a principal fibre bundle, such as $B = M \times G$. Then the lift of ψ to a bundle map $B \to B'$, with $B' = M' \times G$ incorporates the specification of a map $\gamma : M \to G$ that provides the identification of the fibres, i.e., a local gauge transformation. The condition that $\nabla = \psi^*\nabla'$ then means that $\nabla' = \nabla + \gamma^{-1}d\gamma$, and the condition of local covariance of a functional $\mathcal{O}$ now implies that $\mathcal{O}$ can depend on the connection only via its curvature F and its covariant derivatives $\nabla F, \dots, \nabla^{k-2}F$. More generally, if in addition there are dynamical fields Φ valued in an associated bundle $B \times_G V$ (with V a representation of G), then $\mathcal{O}$ can only depend on gauge invariant combinations of $\Phi, \nabla\Phi, \dots, \nabla^k\Phi$. These statements can be proved by the same type arguments as in [76]. For completeness, we give a proof of this generalization of the Thomas replacement theorem incorporating gauge fields in sec. 2.3 below. In our later application to Yang-Mills theory, Φ will consist of Lie-algebra valued vector and ghost fields, in which case V is the Lie-algebra of G, on which G acts via the adjoint representation.

We denote the space of all locally covariant p-form functionals (2) by $\mathbf{P}^p(M)$, or simply by $\mathbf{P}^p$, and we define

$$\mathbf{P}(M) = \bigoplus_{p=0}^n \mathbf{P}^p(M). \tag{5}$$

We also assume for technical reasons that the expressions in $\mathbf{P}$ have at most polynomial dependence upon the dynamical fields Φ, and an analytic dependence upon the background fields Ψ. These definitions can easily be generalized to the case when (Φ, Ψ) are not ordinary fields

valued in some bundle, but instead Grassmann valued fields. A Grassmann valued field is by definition simply a field that is valued in the infinite dimensional exterior algebra E, which is the graded vector space

$$E = \mathrm{Ext}(V) = \bigoplus_n E_n, \quad E_n = \bigwedge^n V \tag{6}$$

with V some infinite-dimensional complex vector space. The space E is equipped with the wedge product $\wedge : E_n \times E_m \to E_{m+n}$, which has the property that $e_n e_m = (-1)^{nm} e_m e_n$ for $e_n \in E_n, e_m \in E_m$, and $e_n e_m = 0$ for all e_n if and only if $e_m = \lambda e_n$. The elements e_n in E_n are assigned Grassmann parity $\varepsilon(e_n) = n$ modulo 2. Thus, when Grassmann valued field are present, expressions $\mathcal{O} \in P^p$ are no longer valued in the p-forms over M, but instead in the set of p-forms over M, tensored with E. A Grassmann valued field consequently has a formal expansion of the form

$$\Phi(x) = \sum_{n \geq 0} e_n \Phi_n(x), \quad e_n \in E_n, \tag{7}$$

where each Φ_n is an ordinary p-form field.

A Lagrangian is a (possibly E-valued) n-form $\mathbf{L} = \mathbf{L}[\Phi, \Psi]$ that is locally and covariantly constructed out of the dynamical fields Φ, the background fields Ψ, and finitely many of its derivatives. For manifolds M carrying an orientation, which we shall assume to be given from now on, one can define a canonical volume n-form $\varepsilon = \varepsilon_{\mu_1 \dots \mu_n} dx^{\mu_1} \wedge \dots \wedge dx^{\mu_n}$ by the standard formula

$$dx = \varepsilon = \sqrt{-g}\, dx^0 \wedge \dots \wedge dx^{n-1} \tag{8}$$

where $x^0, \dots, x^{n-1}$ is right handed, and where $\sqrt{-g}$ is the square root of minus the determinant of $g_{\mu\nu}$. Using the volume n-form, one defines the Hodge dual of a form by

$$* \alpha_{\mu_1 \dots \mu_{n-p}} = \frac{(-1)^p}{(n-p)!} \varepsilon^{\nu_1 \dots \nu_p}{}_{\mu_1 \dots \mu_{n-p}} \alpha_{\nu_1 \dots \nu_p} \tag{9}$$

and it is thereby possible to convert the Lagrangian into a scalar. This is more standard in the physics literature, but for our purposes it will be slightly more advantageous to view $\mathbf{L}$ as an n-form. For compactly supported field configurations, we may form an associated action by integrating the Lagrangian n-form over M,

$$S = \int_M \mathbf{L}\,. \tag{10}$$

We define the left and right variation, $\delta_L S / \delta \Phi(x)$ resp. $\delta_R S / \delta \Phi(x)$ with respect to the dynamical fields by the relation

$$\frac{d}{dt} S[\Phi_t; \Psi] \bigg|_{t=0} = \int_M \frac{\delta_L S}{\delta \Phi(x)} \delta \Phi(x) = \int_M \delta \Phi(x) \frac{\delta_R S}{\delta \Phi(x)}, \quad \delta \Phi(x) = \frac{d}{dt} \Phi_t(x) \bigg|_{t=0}. \tag{11}$$

The left and right derivatives may differ from each other only for Grassmann-valued fields Φ, and we adopt the convention that the left derivative is meant by default if the subscript is suppressed. In terms of the Lagrangian n-form, the variational derivative is given by

$$\frac{\delta S}{\delta \Phi(x)} = \sum_{q=0}^{k} (-1)^q \nabla_{(\mu_1 \dots \mu_q)} \left\{ \frac{\partial \mathbf{L}}{\partial (\nabla_{(\mu_1 \dots \mu_q)} \Phi(x))} \right\} , \tag{12}$$

where we use the abbreviation $\nabla_{(\mu_1 \dots \mu_k)}$ for the k-fold symmetrized derivative in eq. (1). The quantity $\delta S / \delta \Phi(x)$ is an n-form that is locally and covariantly constructed out of the dynamical fields and the background fields and their derivatives, and may hence be viewed as a differential operator acting on Φ. Field configurations Φ satisfying the differential equation

$$\frac{\delta S}{\delta \Phi(x)} = 0 \tag{13}$$

are said to satisfy the equations of motion associated with S, or to be "on shell."

A symmetry is an infinitesimal field variation $s\Phi = \delta\Phi$ of the dynamical fields such that $s\mathbf{L} = d\mathbf{B}$ for some locally constructed $(n-1)$-form $\mathbf{B}$. The existence of symmetries implies the existence of a conserved Noether current, $\mathbf{J}$, defined by

$$\mathbf{J}(\Phi) = \theta(\Phi, s\Phi) - \mathbf{B}(\Phi) , \tag{14}$$

where θ is the $(n-1)$ form defined by

$$\theta_{\nu_1 \dots \nu_{n-1}}(\Phi, \delta\Phi) = \sum_{q=0}^{k-1} \nabla_{(\mu_1 \dots \mu_q)} \delta\Phi \left\{ \frac{\partial \mathbf{L}_{\nu_1 \dots \nu_{n-1} \sigma}}{\partial (\nabla_{(\mu_1 \dots \mu_q \sigma)} \Phi)} \right\} , \tag{15}$$

where we are suppressing the dependence upon the background fields. θ is the boundary term that would arise if $\mathbf{L}$ is varied under an integral sign. As a consequence of the definition, we have

$$d\mathbf{J} = \sum s\Phi_i \frac{\delta S}{\delta \Phi_i} , \tag{16}$$

so $\mathbf{J}$ is indeed conserved on shell. In the context of perturbation theory studied in this paper, the Lagrangian is a power series

$$\mathbf{L} = \mathbf{L}_0 + \lambda \mathbf{L}_1 + \lambda^2 \mathbf{L}_2 + \dots , \tag{17}$$

where $\mathbf{L}_0$ is called the "free Lagrangian" and contains only terms at most quadratic in the dynamical fields Φ, hence giving rise to linear equations of motion. If the symmetry is also a formal power series

$$s = s_0 + \lambda s_1 + \lambda^2 s_2 + \dots , \tag{18}$$

then there is obviously an expansion

$$\mathbf{J} = \mathbf{J}_0 + \lambda \mathbf{J}_1 + \lambda^2 \mathbf{J}_2 + \dots, \tag{19}$$

s_0 is a symmetry of the free Lagrangian $\mathbf{L}_0$ with corresponding conserved Noether current $\mathbf{J}_0$ when the equations of motion hold for $\mathbf{L}_0$.

The theories that we will deal with in this paper all have the property that $\mathbf{L}_0$ contains the highest derivative terms in the dynamical fields $\boldsymbol{\Phi}$. In this case, it is natural to assign a "canonical dimension" to each of the dynamical fields as follows. Let us assume that the background fields consist of a metric, g, and a covariant derivative operator, ∇, which acts like the Levi-Civita connection on tensors. Consider a rescaling of the metric by a constant conformal factor, $\mu^2 g$, where $\mu \in \mathbb{R}$. Then there exists typically a unique rescaling $\boldsymbol{\Phi}_i \to \mu^{d(\Phi_i)} \boldsymbol{\Phi}_i$, $\boldsymbol{\Psi}_i \to \mu^{d(\Psi_i)} \boldsymbol{\Psi}_i$ and $c_i \to \mu^{d(c_i)} c_i$ of the dynamical fields, the background fields, and the coupling constants in $\mathbf{L}_0$ such that $\mathbf{L}_0 \to \mathbf{L}_0$. The numbers $d(\Phi_i), d(\Psi_i)$ and $d(c_i)$ are called the "engineering dimensions" of the fields and the couplings, respectively. The corresponding dimension of composite objects in $\mathbf{P}$ is given by the counting operators $\mathcal{N}_f, \mathcal{N}_c, \mathcal{N}_r : \mathbf{P}(M) \to \mathbf{P}(M)$

$$\begin{aligned} \mathcal{N}_f &= \sum (d(\Phi_i) + k) \, \nabla^k \Phi_i \frac{\partial}{\partial (\nabla^k \Phi_i)} && (20) \\ \mathcal{N}_c &= \sum d(c_i) \, c_i \frac{\partial}{\partial c_i} && (21) \\ \mathcal{N}_r &= \sum (d(\Psi_i) + k) \, \nabla^k \Psi_i \frac{\partial}{\partial (\nabla^k \Psi_i)} \, . && (22) \end{aligned}$$

Not for all S, and not for all choices of the background fields $\boldsymbol{\Psi}$ do the equations of motion (13) possess a well posed initial value formulation, which is a key requirement for a physically reasonable theory. For first order differential equations one can formulate general conditions under which the equations will posses a well-posed initial value formulation. For example, for first order systems of so-called "symmetric hyperbolic type," the initial value problem is well posed in the sense that, given initial data for $\boldsymbol{\Phi}$ on a suitably chosen $n-1$-dimensional hypersurface, there exists a unique solution for sufficiently short "times," i.e., in some open neighborhood of Σ. Furthermore, the propagation of disturbances is "causal" in a well-defined sense, see e.g. [55]. Equations of motion of higher differential order can always be reduced to ones of first order by picking suitable auxiliary field variables, but it is not obvious in a given example which choice will lead to a symmetric hyperbolic system. Fortunately, the equations of motion that we will study in this paper will all be of the form of a simple wave-equation. Actually, since we only consider perturbation theory, we will only be concerned with the existence of solutions for the "free theory," defined by S_0. For the actions considered in this paper, the corresponding equations are linear, and of the form

$$0 = \frac{\delta S_0}{\delta \Phi} = \Box \Phi + (\text{lower order terms}) \tag{23}$$

where $\Box = g^{\mu\nu}\nabla_\mu\nabla_\nu$ is the wave operator in curved space. Such equations do posses a well-posed initial value formulation if the metric does not have any gross causal pathologies, such as closed timelike curves. A typical such equation (for a real scalar field $\Phi = \phi$) is the Klein-Gordon equation

$$(\Box - m^2)\phi = j\,, \tag{24}$$

where m^2 is a constant. For that equation, the initial value problem is well-posed globally for example if the spacetime manifold (M,g) is "globally hyperbolic," meaning by definition that there exists a (necessarily spacelike) "Cauchy-surface", Σ, i.e., a surface which has the property that any inextendible timelike curve hits Σ precisely once. We will always assume in this work that (M,g) is globally hyperbolic. Then, given any f_0, $f_1 \in C_0^\infty(\Sigma)$, there exists a unique solution to eq. (24) such that $\phi|\Sigma = f_0$, and $n^\mu\nabla_\mu\phi|\Sigma = f_1$, where n is the timelike normal to Σ.

The well-posedness of the initial value problem for the Klein-Gordon equation directly leads to the existence of advanced and retarded propagators, which are the uniquely determined distributions Δ_A, Δ_R on $M \times M$ with the properties

$$(\Box - m^2)\Delta_A(x,y) = \delta(x,y) = (\Box - m^2)\Delta_R(x,y) \tag{25}$$

and the support properties

$$\operatorname{supp}\Delta_{A,R} \subset \{(x,y) \in M \times M \mid \quad y \in J^\mp(x)\}\,, \tag{26}$$

where $J^\pm(S)$ denotes the causal future/past of a set $S \subset M$ and is defined as the set of points $x \in M$ with the property that there is a future/past directed timelike or null curve γ connecting x with a point in S.

2.2 Yang-Mills theories, consistency conditions, cohomology

The theory that we are considering in this paper is pure Yang-Mills theory, classically described by the action

$$S_{ym} = -\frac{1}{2}\int_M F^I \wedge *F_I\,. \tag{27}$$

Here, $F_{\mu\nu} = (i/\lambda)[\mathcal{D}_\mu, \mathcal{D}_\nu]$ is the 2-form field strength tensor of a gauge connection $\mathcal{D}$ in some principal G-bundle over M, where G is a direct product of $U(1)^l$ and a semi-simple Lie group. λ is a coupling constant that could be omitted at the classical level. For the sake of simplicity, we will assume that the principal bundle is toplogically trivial, i.e., of the form $M \times G$. We denote the generators of the gauge Lie algebra by $T_I, I = 1,\dots,dim(G)$, and we write $F = T_I F^I_{\mu\nu} dx^\mu \wedge dx^\nu$ for the components of the field strength and similarly for any other Lie-algebra valued field. Lie algebra indices I are raised an lowered with the Cartan-Killing metric k_{IJ} defined by $\mathrm{Tr}\, ad(T_I) ad(T_J)$ for the generators of the semi-simple part, and by 1 for the abelian factors.

The classical field equations for the connection $\mathcal{D}$ derived from is action are

$$g^{\mu\nu}[\mathcal{D}_\mu, [\mathcal{D}_\nu, \mathcal{D}_\sigma]] = 0\,, \tag{28}$$

or, written in more conventional form,

$$\mathcal{D}_{[\mu} * F_{\nu\sigma]} = 0\,. \tag{29}$$

As is particularly clear from the first formulation, the connection $\mathcal{D}$ is the dynamical field variable in this equation. It is convenient to decompose it into a fixed background connection ∇, plus λ times a Lie-algebra valued 1-form field $A = T_I A^I_\mu dx^\mu$,

$$\mathcal{D} = \nabla + i\lambda A\,, \qquad \lambda \in \mathbb{R}\,. \tag{30}$$

The 1-form field A is now the dynamical variable. The coupling constant λ is redundant at the classical level and may be absorbed in A, but it is useful as an explicit perturbation parameter when one wants to study the theory perturbatively. The coupling constant λ acquires a new role at the quantum level due to renormalization effects as we will see below. It is convenient to define ∇ on tensor fields to be the standard Levi-Civita connection of the metric. The background derivative operator then has the curvature tensor

$$[\nabla_\mu, \nabla_\nu] k_\sigma = R_{\mu\nu\sigma}{}^\rho k_\rho + f^I_{\mu\nu} \mathrm{R}(T_I) k_\sigma \tag{31}$$

where R is the representation of the Lie-algebra associated with k_μ, and $f = T_I f^I_{\mu\nu} dx^\mu \wedge dx^\nu$ is the curvature of the background gauge connection. In Minkowski space, it is typically assumed that $\nabla = \partial$, implying that $f = 0$. For simplicity, we will assume that the background gauge connection has been chosen as the standard flat connection in the bundle $M \times G$, so that $f = 0$ on our manifold M. The advantage of this choice is that all quantities that are locally and covariantly constructed out of the field A and the background structure $g = g_{\mu\nu} dx^\mu dx^\nu$ and ∇ can be written in the form eq. (4) with $\Phi = A$, without any explicit appearance of the background curvature f. If $f \neq 0$, we would have to include everywhere explicitly the background curvature.

With the decomposition $\mathcal{D} = \nabla + i\lambda A$, the curvature F is given by

$$F^I_{\mu\nu} = \nabla_\mu A^I_\nu - \nabla_\nu A^I_\mu + i\lambda f^I{}_{JK} A^J_\mu A^K_\nu \tag{32}$$

where $f^I{}_{JK}$ are the structure constants of the Lie-algebra defined by $[T_I, T_J] = f_{IJ}{}^K T_K$. The equations of motion, when written in terms of A, are not hyperbolic, in the sense that the highest derivative term is not of the form of a wave equation. Thus, the equations of motion for Yang-Mills theory do not straightforwardly admit an initial value formulation. This feature is a consequence of the fact that the Yang-Mills Lagrangian and equations of motion is invariant under the group of local gauge transformations acting on the dynamical fields by $\mathcal{D} \mapsto \gamma(x)^{-1} \mathcal{D} \gamma(x)$, where $g : M \to G$ is any smooth function valued in the group, or equivalently by

$$\nabla \mapsto \nabla\,, \quad A \mapsto \gamma^{-1} A \gamma - \frac{i}{\lambda} \gamma^{-1} d\gamma. \tag{33}$$

Since such local gauge transformations allow one to make local changes to the dynamical field variables, it is clear that those are not entirely specified by initial conditions. However, the freedom of making local gauge transformation can be used to set some components of A to zero, so that the remaining components satisfy a hyperbolic equation and consequently admit a well-posed initial value formulation, as described e.g. in [1]. Later, we want to perturbatively construct a quantum version of Yang-Mills theory, and for this purpose, another approach seems to be much more convenient. This approach consists in adding further fields to the theory which render the equations of motion hyperbolic, and which can, at a final stage, be removed by a symmetry called "BRST-symmetry".

In the BRST approach, one introduces additional dynamical Grassmann Lie-algebra valued fields $C = T_I C^I, \bar{C} = T_I \bar{C}^I$, and a Lie-algebra valued field $B = T_I B^I$, and one defines a new theory with action S_{tot} by

$$S_{tot} = S_{ym} + S_{gf} + S_{gh}\,, \tag{34}$$

where S_{gf} is a "gauge fixing" term defined by

$$S_{gf} = \int_M B^I (i\mathcal{G}_I + \frac{1}{2}B_I) \tag{35}$$

with a local covariant "gauge fixing" functional $\mathcal{G}$ of the field A, and where S_{gh} is the "ghost" term, defined by

$$S_{gh} = i\int_M \mathcal{D}_\mu C^J \frac{\delta(\mathcal{G}^I \bar{C}_I)}{\delta A^J_\mu}\,\varepsilon\,. \tag{36}$$

The total set of dynamical fields is denoted $\Phi = (A^I, C^I, \bar{C}^I, B^I)$, and their assignment of ghost number, Grassmann parity, dimension, and form degree is summarized in the following table

Φ	A^I	C^I	$\bar{C}^I$	B^I
Dimension	1	0	2	2
Ghost Number	0	1	-1	0
Form Degree	1	0	0	0
Grassman Parity	0	1	1	0

The assignments of the dimensions are given for the case when the spacetime M is 4-dimensional, to which we now restrict attention for definiteness. To state the relation between the auxiliary theory and the original Yang-Mills theory, one first observes that the action S_{tot} of the auxiliary theory is invariant under the following so-called BRST-transformations [9, 10]:

$$\begin{aligned} sA^I &= dC^I + i\lambda f^I{}_{JK} A^J C^K, & (37)\\ sC^I &= -\frac{i\lambda}{2} f^I{}_{JK} C^I C^K, & (38)\\ s\bar{C}^I &= B^I\,, & (39)\\ sB^I &= 0\,. & (40) \end{aligned}$$

The assignment of the various numbers to the fields are done in such a way that s has dimension 0, ghost number +1, grassmann parity +1, and form degree 0. It is declared on arbitrary local covariant functionals $\mathcal{O} \in \mathbf{P}(M)$ of the dynamical fields $A, C, \bar{C}, B$ and the background fields by the rules $\nabla \circ s - s \circ \nabla = 0 = dx^\mu \circ s + s \circ dx^\mu$, and on (wedge) products via the graded Leibniz rule, $s(\mathcal{O}_p \wedge \mathcal{O}_q) = s\mathcal{O}_p \wedge \mathcal{O}_q + (-1)^{p+\varepsilon(\mathcal{O}_p)} \mathcal{O}_p \wedge s\mathcal{O}_q$. With these definitions, it follows that

$$s^2 = 0\,, \quad sd + ds = 0\,. \tag{41}$$

The key equation is

$$sS_{tot} = 0\,, \tag{42}$$

which one may verify by writing S_{tot} in the form

$$S_{tot} = S_{ym} + s\Psi \tag{43}$$

where

$$\Psi = \int_M \bar{C}^I (\frac{1}{2} B_I + i\mathcal{G}_I)\varepsilon\,. \tag{44}$$

Indeed, S_{ym} is invariant because s just acts like an ordinary infinitesimal gauge transformation on A, while s annihilates the second term because $s^2 = 0$. In this paper, we choose the gauge fixing functional as

$$\mathcal{G}^I = \nabla^\mu A^I_\mu\,. \tag{45}$$

Then the equation of motion for B^I is algebraic, $B^I = -i\nabla^\mu A^I_\mu$. Inserting this into the equation of motion for A^I_μ, one sees that this equation is of the form (23). Indeed, this special choice of the gauge fixing function effectively eliminates a term of the form $\nabla^\mu \nabla_\nu A^I_\mu$ (which would spoil hyperbolicity) from the equations of motion for the gauge field, thus leaving only the wave operator. The remaining equations for $C^I, \bar{C}^I$ are also of the form (23). Thus, the equations of motion for the total action S_{tot} are of wave equation type. They consequently possess a well-posed initial value formulation at the linear level, which is sufficient for perturbation theory, and in fact also at the non-linear level [1].

Given that S_{tot} defines a classical theory with a well-posed initial value formulation, we may define an associated graded Peierls bracket [35, 34, 91, 27, 87] $\{\mathcal{O}_1, \mathcal{O}_2\}_{\text{P.B.}}$, for any pair of local[3] functionals $\mathcal{O}_1, \mathcal{O}_2 \in \mathbf{P}$. Since the action S_{tot} is invariant under s, it follows that the (graded) Peierls bracket is also invariant under s, in the sense that

$$s\{\mathcal{O}_1, \mathcal{O}_2\}_{\text{P.B.}} = \{s\mathcal{O}_1, \mathcal{O}_2\}_{\text{P.B.}} + (-1)^{\varepsilon(\mathcal{O}_1)+deg(\mathcal{O}_1)}\{\mathcal{O}_1, s\mathcal{O}_2\}_{\text{P.B.}}\,, \tag{46}$$

$(-1)^{\varepsilon(\mathcal{O}_1)}$ denoting the Grassmann parity of a functional of the fields, and $deg(\mathcal{O}_1)$ the form degree. The connection between the classical auxiliary theory associated with S_{tot}, and Yang-Mills theory with action S_{ym} is based on the following key Lemma:

[3]The Peierls bracket may also be defined for certain non-local functionals. The consideration of such functionals is necessary in order to contain a set of functionals that is stable under the bracket.

Lemma 1. Let $\mathcal{O} \in \mathbf{P}$ be a local covariant functional of the background connection, the background metric, and the fields $\Phi = (A, C, \bar{C}, B)$. Let $s\mathcal{O} = 0$. Then, up to a term of the form $s\mathcal{O}'$, $\mathcal{O}$ is a linear combination of elements of the form

$$\mathcal{O} = \prod_k r_{t_k}(g, R, \nabla R, \dots, \nabla^k R) \prod_i p_{r_i}(C) \prod_j \Theta_{r_j}(F, \mathcal{D}F, \dots, \mathcal{D}^l F), \tag{47}$$

where p_r, Θ_s are invariant polynomials of the Lie-algebra of G, where $F = F_{\mu\nu} dx^\mu \wedge dx^\nu$, and where r_t is a local functional of the metric g, and the Riemann tensor R and its derivatives.

The lemma is essentially a standard result in BRST-cohomolgy, see e.g. [5] and the references cited there. The only difference to the formulation given in [5] is that, in the present setting, the coefficients r_t can only depend locally and covariantly upon the metric (as opposed to being an arbitrary form on spacetime). The fact that r_t then has to be a functional of the Riemann tensor and its derivatives follows again from the "Thomas replacement argument", see e.g. [76] and the next subsection. Thus, at zero ghost number, the local and covariant functionals in the kernel of s are precisely the local gauge invariant observables of Yang-Mills theory modulo an element in the image of s, so the equivalence classes of the kernel of s modulo the image of s at zero ghost number,

$$\{\text{class. gauge. inv. fields}\} = \frac{\text{Kernel}\, s}{\text{Image}\, s} \quad (\text{at zero ghost number}), \tag{48}$$

are in one-to-one correspondence with the gauge invariant observables. Furthermore, by (46), the brackets are well-defined on the cohomology classes, and the Yang-Mills equations of motion hold modulo s. Thus, the theory whose observables are defined by the equivalence classes of s (at zero ghost number), and whose bracket is defined by the Peierls bracket may be viewed as a definition of classical Yang-Mills theory.

The BRST-transformation s plays a crucial role also in the perturbative quantum field theory associated with Yang-Mills theory, where its role is among other things to derive certain consistency conditions on the terms in the renormalized perturbation series. We therefore discuss some of the relevant facts about the BRST-transformation in some more detail. Since $s^2 = 0$, the BRST transformation defines a "differential", or, more precisely, a differential complex

$$s : \mathbf{P}_0 \to \mathbf{P}_1 \to \cdots \to \mathbf{P}_N \to \dots \tag{49}$$

where a subscript denotes the grading of the functionals in $\mathbf{P}$ by the ghost number, defined by the ghost number operator $\mathcal{N}_g$ counting the ghost number of an element in $\mathbf{P}$ by the formula

$$\mathcal{N}_g = \sum \nabla^k C^I \frac{\partial}{\partial(\nabla^k C^I)} - \nabla^k \bar{C}^I \frac{\partial}{\partial(\nabla^k C^I)}\,. \tag{50}$$

Thus, $\mathbf{P}$ is doubly graded space, by the form degree and ghost number, and we write $\mathbf{P}^q_p$ for the subspace of elements with form degree q and ghost number p. We define the cohomology

ring $H^p(s, \mathbf{P}^q)$ to be the set of all local covariant q-form functionals $\mathcal{O}$ of ghost number p, and $s\mathcal{O} = 0$, modulo the set of a q-form functionals $\mathcal{O} = s\mathcal{O}'$ with ghost number p, i.e.,

$$H^p(s, \mathbf{P}^q(M)) = \frac{\{\text{Kernel}\, s : \mathbf{P}^q_p \to \mathbf{P}^q_{p+1}\}}{\{\text{Image}\, s : \mathbf{P}^q_{p-1} \to \mathbf{P}^q_p\}}\,. \tag{51}$$

The above lemma may be viewed as the determination of the space $H^q(s, \mathbf{P}^p)$ for all q, p. We will also encounter another cohomology ring, consisting of all s-closed local covariant functionals modulo exact local covariant functionals. To describe this ring more precisely, it is useful to know the following result, sometimes called "algebraic Poincare Lemma", or "fundamental Lemma of the calculus of variations":

Lemma 2. (Algebraic Poincare lemma) Let $\alpha = \alpha[\Phi, \Psi]$ be a p-form on an n-dimensional manifold M, which is locally and covariantly constructed out of a number of dynamical fields Φ, and background fields Ψ. Assume that $d\alpha[\Phi, \Psi] = 0$ for all Ψ, and that each Ψ is pathwise connected to a reference Ψ_0 for which $\alpha[\Phi, \Psi_0] = 0$. Then $\alpha = d\beta$ for some $\beta = \beta[\Phi, \Psi]$ which is locally constructed out of the fields.

The proof is given for convenience in the next subsection. Consider now a $\mathcal{O}_q \in \mathbf{P}^q$ such that $s\mathcal{O}_q = d\mathcal{O}_{q-1}$, i.e., $\mathcal{O}_q$ is s-closed modulo d. Then, by $s^2 = 0$ and $ds + sd = 0$, the form $s\mathcal{O}_{q-1}$ is d-closed, and hence d-exact by the fundamental lemma, so $s\mathcal{O}_{q-1} = d\mathcal{O}_{q-2}$. We can now repeat this procedure until we have reached the forms of degree 0, thereby arriving at what is called a "decent-equation", or a "ladder":

$$\begin{aligned} s\mathcal{O}_q &= d\mathcal{O}_{q-1} && (52)\\ s\mathcal{O}_{q-1} &= d\mathcal{O}_{q-2} && (53)\\ &\ \ \dots && (54)\\ s\mathcal{O}_1 &= d\mathcal{O}_0 && (55)\\ s\mathcal{O}_0 &= 0\,. && (56) \end{aligned}$$

Note that, within each ladder, the form degree plus the ghost number is constant. We denote the space of $\mathcal{O}_q$ that are s-closed modulo d at ghost number p, factored by elements that are s-exact modulo d by $H^p(s|d, \mathbf{P}^q) \equiv H^p(s, H^q(d, \mathbf{P}))$. In practice, ladders can be used to determine the cohomology of s modulo d.

For the purpose of perturbative quantum field theory, it will be convenient to consider yet another cohomology ring related to s that incorporates also the equations of motion. Let us add to the theory a further set of background fields ("BRST sources," or "anti-fields" [6]) $\Phi^\ddagger = (A^\ddagger_I, C^\ddagger_I, \bar{C}^\ddagger_I, B^\ddagger_I)$ corresponding to the dynamical fields $\Phi = (A^I, C^I, \bar{C}^I, B^I)$:

$\Phi^\ddagger$	$A^\ddagger_I$	$C^\ddagger_I$	$\bar{C}^\ddagger_I$	$B^\ddagger_I$
Dimension	3	4	2	2
Ghost Number	-1	-2	0	-1
Form Degree	3	4	4	4
Grassmann Parity	1	0	0	1

Consider now the action

$$S[\Phi,\Phi^{\ddagger}] = S_{ym}+S_{gf}+S_{gh}+S_{sc}, \quad S_{sc} = \int_M s\Phi_i \wedge \Phi^{\ddagger i} \tag{57}$$

The new action is still BRST-closed, $sS = 0$, because it is given by the sum of S_{tot} and a BRST-exact term, and it satisfies in addition $(S,S) = 0$, where the "anti-bracket" $(\,.\,,\,.\,)$ is defined by the equation [6]

$$(F_1,F_2) = \int_M \left[\frac{\delta_R F_1}{\delta\Phi_i(x)} \wedge \frac{\delta_L F_2}{\delta\Phi^{\ddagger i}(x)} - (-1)^{deg(\Phi_i^{\ddagger})} \frac{\delta_R F_1}{\delta\Phi^{\ddagger i}(x)} \wedge \frac{\delta_L F_2}{\delta\Phi_i(x)} \right]. \tag{58}$$

The local anti-bracket satisfies the graded Jacobi-identity

$$(-1)^{\varepsilon_3\varepsilon_1}((F_1,F_2),F_3)+(-1)^{\varepsilon_2\varepsilon_1}((F_2,F_3),F_1)+(-1)^{\varepsilon_3\varepsilon_2}((F_3,F_1),F_2) = 0 \tag{59}$$

and as a consequence $(F,(F,F)) = 0$ for any F. The differential incorporating the equations of motion is defined by

$$\hat{s}F = (S,F)\,. \tag{60}$$

It satisfies $\hat{s}^2 = 0$ as a consequence of $(S,S) = 0$ and the Jacobi identity, as well as, $\hat{s}d + d\hat{s} = 0$. It differs from the BRST-differential s by the "Koszul-Tate-differential" σ

$$\hat{s} = s + \sigma\,, \tag{61}$$

where $\sigma^2 = 0$, and σ is anticommuting with s. It acts on the fields by

$$\sigma\Phi_i = 0, \quad \sigma\Phi^{\ddagger i} = \frac{\delta S}{\delta\Phi_i}\,. \tag{62}$$

Thus, acting with σ on a monomial in $\mathbf{P}$ containing an anti-field automatically gives an expression containing a factor of the equations of motion, i.e., an on-shell quantity. This will be useful in the context of perturbative quantum field theory in order to keep track of such terms. Starting from the differential $\hat{s}$, one can again define cohomology rings $H^p(\hat{s},\mathbf{P}^q)$ and $H^p(\hat{s}|d,\mathbf{P}^q)$. The ring $H^0(\hat{s},\mathbf{P}^q)$ is still described by Lemma 2, because one can prove in general that $H^p(s,\mathbf{P}^q)$ and $H^p(\hat{s},\mathbf{P}^q)$ are isomorphic, see e.g. [5]. The relative cohomology rings $H^p(\hat{s}|d,\mathbf{P}^q)$ appear in the analysis of gauge invariance in quantum Yang-Mills theory. They are also known, but they depend somewhat upon the choice of the gauge group G. They are described by the following theorem, see e.g. [5]:

Theorem 1. Let the Lie-group G be semi-simple with no abelian factors, and let $n = dim(M)$. Then each class in $H(\hat{s}|d,\mathbf{P}^n)$ is a linear combination of expressions $\mathcal{O}$ of the form (47), and representatives $\mathcal{O}'$ of the form

$$\mathcal{O}' = n\text{-form part of} \quad \prod_k r_{t_k}(R,\nabla R,\ldots,\nabla^{n_k}R)\prod_i q_{r_i}(C+A,F)\prod_j f_{s_j}(F)\,, \tag{63}$$

where $q_{r_i}(A+C,F)$ are the Chern-Simons forms,

$$q_r(A+C,F) = \int_0^1 \mathrm{Tr}\Big((C+A)[tF+\lambda t(t-1)(C+A)^2]^{m(r)-1}\Big)\,dt\,. \tag{64}$$

where f_s are strictly gauge-invariant monomials of F containing only the curvature, F, but not its derivatives. The numbers $m(r)$ are the degrees of the independent Casimir elements of G, and the trace is in some representation. The r_t are taken to be a basis of closed forms $dr_t = 0$ that are analytic functions of the metric and the covariant derivatives of the Riemann tensor. For $p < n$, a basis of $H(\hat{s}|d,\mathbf{P}^p)$ is given by the $\mathcal{O}'$ at form degree p, together with all elements $\mathcal{O}$ of the form (47), for any Lie-group $H = U(1)^l \times G$, with G semi-simple.

Remarks: 1) The statement of the theorem given in [5] only asserts that the r_t are closed forms on M. To obtain that the r_t in fact have to be analytic functions of $R, \nabla R, \nabla^2 R, \dots$, one has to use that, as we are assuming, the elements in $\mathbf{P}^q$ are locally and covariantly constructed out of the metric in the sense described above, with an analytic dependence upon the spacetime metric. It then follows from the "Thomas replacement argument" [76] that the r_t have to be analytic functions of the curvature tensor and its derivatives. It furthermore follows that the r_t may be chosen to be characteristic classes

$$r_t = \mathrm{Tr}\big(R\wedge\cdots\wedge R\big), \tag{65}$$

where Tr is the trace in a representation of the Lie-algebra of $SO(n-1,1)$, and where $R = T_{ab}R^{ab}_{\mu\nu}dx^\mu \wedge dx^\nu$ is the curvature 2-form of the metric, identified with a 2-form valued in the Lie-algebra of $SO(n-1,1)$ via a tetrad field $e^a_\mu dx^\mu$.

2) There are more elements in $H(\hat{s}|d,\mathbf{P}^n)$ when the group G has abelian factors, see e.g. for a discussion [60]. In pure Yang-Mills theory, abelian factors decouple and hence can be treated separately.

In perturbation theory, we expand S as

$$S = S_0 + \lambda S_1 + \lambda^2 S_2\,, \tag{66}$$

and we correspondingly expand the Lagrangian as

$$\begin{aligned}
\mathbf{L}_0 &= \frac{1}{2}dA^I\wedge *dA_I - id\bar{C}^I\wedge *dC_I + B^I(id*A_I + \frac{1}{2}*B_I) + s_0 A_I\wedge A^{\ddagger I}\,, && (67)\\
\mathbf{L}_1 &= \frac{1}{2}f_{IJK}*dA^I\wedge A^J\wedge A^K + f_{IJK}\bar{C}^I\wedge A^J\wedge *dC^K \\
&\quad + s_1 A_I\wedge A^{\ddagger I} + s_1 C_I\wedge C^{\ddagger I} + s_1\bar{C}_I\wedge \bar{C}^{\ddagger I} && (68)\\
\mathbf{L}_2 &= \frac{1}{4}f^I{}_{JK}f_{ILM}A^J\wedge A^K*(A^L\wedge A^M) && (69)
\end{aligned}$$

in our choice of gauge (45). We correspondingly have an expansion of the Slavnov Taylor differential as $\hat{s} = \hat{s}_0 + \lambda\hat{s}_1 + \lambda^2\hat{s}_2$, and similarly of the Koszul Tate differential as $\sigma = \sigma_0 +$

$\lambda\sigma_1 + \lambda^2\sigma_2$. The zeroth order parts of these expansions still define differentials. The free Slavnov Taylor differential $\hat{s}_0 \mathcal{O} = (S_0, \mathcal{O})$, decomposed as

$$\hat{s}_0 = s_0 + \sigma_0 \tag{70}$$

will play an important role in perturbative quantum field theory. Its action is given explicitly by

$$\hat{s}_0 A^I = dC^I, \quad \hat{s}_0 C^I = 0, \quad \hat{s}_0 \bar{C}^I = B^I, \quad \hat{s}_0 B^I = 0 \tag{71}$$

on the fields, where it coincides with that of s_0. Its action on the anti-fields is given by

$$\hat{s}_0 A_I^\ddagger = \frac{\delta S_0}{\delta A^I}, \quad \hat{s}_0 C_I^\ddagger = \frac{\delta S_0}{\delta C^I}, \quad \hat{s}_0 \bar{C}_I^\ddagger = \frac{\delta S_0}{\delta \bar{C}^I}, \quad \hat{s}_0 B_I^\ddagger = \frac{\delta S_0}{\delta B^I} \tag{72}$$

where it coincides with that of σ_0. The actions of s_0 and σ_0 are summarized in the following table:

Field	s_0	σ_0
A^I	dC^I	0
B^I	0	0
C^I	0	0
$\bar{C}^I$	B^I	0
$A_I^\ddagger$	0	$-d * dA_I - i * dB_I$
$B_I^\ddagger$	0	$-id * A_I$
$C_I^\ddagger$	0	$id * d\bar{C}_I - dA_I^\ddagger$
$\bar{C}_I^\ddagger$	0	$id * dC_I$

In perturbation theory, if $F = F_0 + \lambda F_1 + \lambda^2 F_2 + \dots$, equations like $sF = 0$ are understood in terms perturbative sense, as the hierarchy of identities obtained by expanding the terms out in λ. This makes no difference with regard to the above 2 cohomological lemmas, which now also have to be interpreted in the sense of formal power series (in fact, the proof of those lemmas is perturbative). We finally metion a few identities satisfied by the BRST-current $\mathbf{J}$ defined above that we will need later. First, from the expression for the differential of the BRST current, we have

$$d\mathbf{J}(x) = \sum_i (S, \Phi_i(x))(\Phi^{\ddagger i}(x), S)\,. \tag{73}$$

Applying the differential $\widehat{s} = (S, .)$ and using the Jacobi identity for the anti-bracket as well as $(S, S) = 0$, we get

$$d\hat{s}\mathbf{J} = 0\,, \tag{74}$$

so by lemma 3, we have the identity

$$\hat{s}\mathbf{J} = d\mathbf{K}\,, \tag{75}$$

for some $(n-2)$-form $\mathbf{K}$, which is the beginning of a cohomologically trivial ladder, i.e., $\mathbf{J}$ is the zero element in $H^1(\hat{s}|d,\mathbf{P}^3)$. If we expand this identity in λ, we get

$$\hat{s}_0\mathbf{J}_0 = d\mathbf{K}_0, \quad \hat{s}_1\mathbf{J}_0 + \hat{s}_0\mathbf{J}_1 = d\mathbf{K}_1, \quad \text{etc.} \tag{76}$$

The free BRST-current $\mathbf{J}_0$ and its perturbation $\mathbf{J}_1$ are given by

$$\begin{aligned}
\mathbf{J}_0 &= *dA^I \wedge dC_I - iB^I * dC_I = \hat{s}_0(dA^I \wedge A_I - i\bar{C}^I \wedge *dC_I)\,, \qquad (77)\\
\mathbf{J}_1 &= i f_{IJK}[C^J A^K \wedge *A^I + \frac{i}{2} * d\bar{C}^I C^J C^K + iB^I C^J * A^K - \frac{1}{2} C^J C^K A^{\ddagger I} + \frac{1}{2} dC^I \wedge *(A^J \wedge A^K)]\,,
\end{aligned}$$

and the $\mathbf{K}_i$ are given by $\mathbf{K}_0 = 0$, and

$$\mathbf{K}_1 = \frac{-i}{2} f_{IJK} dA^I C^J C^K \tag{78}$$

$$\mathbf{K}_2 = \frac{1}{2} f_{IJK} f^I{}_{MN} A^I \wedge A^K C^M C^N\,. \tag{79}$$

Another fact that we will need later is the equation

$$\hat{s}\mathbf{L} = d\mathbf{J}\,, \tag{80}$$

where $\psi = (\Phi,\Phi^\ddagger)$ collectively denotes the fields and anti-fields, and where we recall that $\mathbf{J}[\psi] = \theta[\psi, s\Phi]$, where θ was defined in eq. (15). To prove this relation, we recall that $\hat{s} = s + \sigma$, so

$$\hat{s}\mathbf{L}[\psi] = \sigma\mathbf{L}[\psi] + \frac{\delta S[\psi]}{\delta\Phi} \wedge s\Phi + d\theta[\psi, s\Phi] = \sigma\mathbf{L}[\psi] + \sigma\Phi^\ddagger \wedge s\Phi + d\theta[\psi, s\Phi] = d\theta[\psi, s\Phi]\,, \tag{81}$$

using in the last line that $\sigma\mathbf{L} = -\delta S/\delta\Phi \wedge s\Phi$. The equation we have just derived may be expanded in powers of λ, leading to the relations

$$\hat{s}_0\mathbf{L}_0 = d\mathbf{J}_0\,, \tag{82}$$

$$\hat{s}_0\mathbf{L}_1 + \hat{s}_1\mathbf{L}_0 = d\mathbf{J}_1\,. \tag{83}$$

2.3 Proof of the Algebraic Poincare Lemma, and the Thomas Replacement Theorem

Lemma 3. (Algebraic Poincare Lemma) Let $\alpha = \alpha[\Phi,\Psi]$ be a p-form on an n-dimensional manifold M, which is locally and covariantly constructed out of a number of dynamical fields Φ, and background fields Ψ. Assume that $d\alpha[\Phi,\Psi] = 0$ for all Ψ, and that each Ψ is pathwise connected to a reference Ψ_0 for which $\alpha[\Phi,\Psi_0] = 0$. Then $\alpha = d\beta$ for some $\beta = \beta[\Phi,\Psi]$ which is locally constructed out of the fields.

The algebraic Poincare lemma has been rediscovered many times, and different proofs exist in the literature. Here we follow the proof given in [109], for other accounts see e.g. [5].
Proof: One first considers the case when $\alpha[\Phi,\Psi]$ is linear in Ψ, i.e., of the form

$$\alpha_{\mu_1\dots\mu_p} = \sum_{i=0}^{k} A^i{}_{\mu_1\dots\mu_p}{}^{\nu_1\dots\nu_i}(\Phi)\nabla_{(\nu_1}\cdots\nabla_{\nu_i)}\Psi\,, \tag{84}$$

where we may assume that A^i is totally symmetric in the upper indices, and totally anti-symmetric in the lower indices. The condition that $d\alpha = 0$ implies the condition

$$A^k{}_{[\mu_1\dots\mu_p}{}^{\nu_1\dots\nu_k}\delta_{\gamma]}{}^{\delta}\nabla_{(\delta}\nabla_{\nu_1}\dots\nabla_{\nu_k)}\Psi = 0\,. \tag{85}$$

At each $x \in M$, $\nabla_{(\nu_1}\dots\nabla_{\nu_k)}\Psi|_x$ can be chosen to be an arbitrary totally symmetric tensor, so we must have

$$A^k{}_{[\mu_1\dots\mu_p}{}^{(\nu_1\dots\nu_k}\delta_{\gamma]}{}^{\delta)} = 0\,. \tag{86}$$

Contracting over δ,γ and using the symmetries of A^k, one finds

$$\begin{aligned}&\Big[\frac{n}{(k+1)(p+1)} + \frac{k}{(k+1)(p+1)} - \frac{p}{(k+1)(p+1)}\Big]A^k{}_{\mu_1\dots\mu_p}{}^{\nu_1\dots\nu_k} - \\ &-\frac{kp}{(k+1)(p+1)}A^k{}_{\gamma[\mu_2\dots\mu_p}{}^{\gamma(\nu_2\dots\nu_k}\delta_{\mu_1]}{}^{\nu_1)} = 0\end{aligned} \tag{87}$$

and therefore that

$$A^k{}_{\mu_1\dots\mu_p}{}^{\nu_1\dots\nu_k} = \frac{kp}{(k+1)(p+1)}A^k{}_{\gamma[\mu_2\dots\mu_p}{}^{\gamma(\nu_2\dots\nu_k}\delta_{\mu_1]}{}^{\nu_1)}\,. \tag{88}$$

For $k = 0$, this condition simply reduces to $A^0 = 0$ and hence $\alpha = 0$, thus proving that the lemma is trivially fulfilled when $k = 0$ and when α depends linearly on Ψ. For $k > 0$, one may proceed inductively. Thus, assume that the statement has been shown for all $k \leq m-1$. Define

$$\tau_{\mu_2\dots\mu_m} = \frac{mp}{(m+1)(p+1)}A^m{}_{\gamma[\mu_2\dots\mu_p]}{}^{\gamma\nu_2\dots\nu_m}\nabla_{(\nu_2}\cdots\nabla_{\nu_m)}\Psi\,, \tag{89}$$

and let

$$\alpha' = \alpha - d\tau\,. \tag{90}$$

Then α' is still closed and locally constructed from Φ,Ψ, linear in Ψ, but by (194), it only contains terms with a maximum number $m-1$ of derivatives on Φ. For such α', we inductively know that $\alpha' = d\gamma$ for a locally constructed γ. Thus, $\alpha = d(\gamma+\tau)$, thereby closing the induction loop. Thus, we have proved the lemma when α depends linearly upon Ψ.

Consider now the case when $\alpha[\Phi,\Psi]$ is non-linear in Ψ. Let $\tau \mapsto \Psi_\tau$ be a smooth path in field space with $\Psi_0 = \Psi$. Putting $\frac{d}{d\tau}\Psi|_{\tau=0} = \delta\Psi$, we have

$$d\Big\{\frac{d}{d\tau}\alpha[\Phi,\Psi_\tau]\big|_{\tau=0}\Big\} = d\Big\{\sum_{i=1}^{k}\frac{\partial\alpha[\Phi,\Psi]}{\partial(\nabla_{(\mu_1}\dots\nabla_{\mu_i)}\Psi)}\,\nabla_{(\mu_1}\dots\nabla_{\mu_i)}\delta\Psi\Big\} = 0\,. \tag{91}$$

Since this must hold for all paths, the identity holds for all $\Phi, \Psi, \delta\Psi$. Thus, since this expression is linear in $\delta\Psi$ and must hold for all $\delta\Psi$, we can find a γ such that

$$\frac{d}{d\tau}\alpha[\Phi, \Psi_\tau]\bigg|_{\tau=0} = d\gamma[\Phi, \Psi, \delta\Psi]\,. \tag{92}$$

where γ is constructed locally out of the fields. Thus, for any path in field space, we have

$$\alpha[\Phi, \Psi_\tau] = \alpha[\Phi, \Psi_0] + d\left\{ \int_0^\tau \gamma\Big[\Phi, \Psi_t, \frac{d}{dt}\Psi_t\Big]\, dt \right\}. \tag{93}$$

Consequently, for any field configuration Ψ that can be reached by a differentiable path from a reference configuration Ψ_0 for which $\alpha(\Phi, \Psi_0)$, we can write $\alpha(\Phi, \Psi) = d\beta(\Phi, \Psi)$. $\square$

We next give the precise statement and proof of the Thomas replacement theorem in the case that we have gauge fields, a metric and background fields. We consider spacetime manifolds (M, g), and G principal fibre bundles $B \to M$ over M with an arbitrary but fixed structure group G. On B, we consider gauge connections $\mathcal{D}$. As above, if we have any section k in $(T^*M)^{\otimes m} \otimes (TM)^{\otimes n}$ times $B \times_G V$, where V is a vector space with an action of G, then we let $\mathcal{D}$ act on the "tensor part" of k by the Levi-Civita connection ∇ of g, and on the "fibre bundle part" by $\mathcal{D}$. We denote by $j^p_x(g, \mathcal{D}, \Psi)$ the p-jet of the metric and the gauge connection and background field, and we consider functionals

$$\mathcal{O}(x) = \mathcal{O}\,[\, j^p_x(g, \mathcal{D}, \Psi)]\,. \tag{94}$$

Let $\psi : B \to B$ be a bundle morphism, i.e., a diffeomorphism of B which is compatible with the G-action on B in the sense that $g\psi(y) = \psi(gy)$. Let $g, \mathcal{D}$ be a metric and connection on B, and let $\psi^* g, \psi^* \mathcal{D}$ be the pull-backs. If $\mathcal{O}$ depends locally and covariantly upon the metric and connection, then we have

$$\psi^* \mathcal{O}\,[\, j^p(\psi^* g, \psi^* \mathcal{D}, \Psi)] = \mathcal{O}\,[\, j^p(\psi^* g, \psi^* \mathcal{D}, \Psi)]\,, \tag{95}$$

where we note that ψ^* does *not* act on the background fields on the right side. This equation is to hold for *all* $g, \mathcal{D}$, and *some* choice of the background field(s) Ψ. If $B = M \times G$, then the above condition can be stated somewhat more explicitly as follows. We may identify the p-jet of the background fields Ψ with a collection of tensor fields on M (including the derivatives of any background field), which we again denote by Ψ for simplicity. Let us introduce an arbitrary *background* derivative operator D (no relation to $\mathcal{D}$ is assumed), and consider first the case when ψ is a "pure diffeomorphism", i.e., $\psi = f \times id_G$, with f a diffeomorphism of M. Let us decompose $\mathcal{D}$ as $\mathcal{D} = D + i\lambda A$, with A a Lie-algebra valued 1-form on M. Then the above condition can be written as

$$f^* \mathcal{O}\,[g, \dots, D^p g, A, \dots, D^p A, \Psi] = \mathcal{O}\,[\, f^* g, \dots, D^p f^* g, f^* A, \dots, D^p f^* A, \Psi]\,. \tag{96}$$

where as usual we denote by $D^k = dx^{\mu_1} \otimes \dots dx^{\mu_k} D_{(\mu_1} \dots D_{\mu_k)}$ the symmetrized k-fold derivative. Note that, in the above expression, f^* does not act on any of the background fields Ψ, nor on D.

(The background fields Ψ may in particular include the curvature of the background connection D, and their derivatives are regarded as "independent fields".) Secondly, let ψ be a "pure gauge transformation", i.e., a transformation of the form $\psi = id_M \times \gamma$, where $\gamma : M \to G$ is a local gauge transformation. Let $A^\gamma = \gamma^{-1} A \gamma - (i/\lambda)\, \gamma^{-1} d\gamma$. Then the above condition (95) becomes

$$\mathcal{O}[g, \dots, D^p g, A, \dots, D^p A, \Psi] = \mathcal{O}[g, \dots, D^p g, A^\gamma, \dots, D^p A^\gamma, \Psi]\,. \tag{97}$$

Lemma 4. (Thomas Replacement Theorem) If $\mathcal{O}$ is a functional satisfying eq. (95) [or equivalently eqs. (96) and (97) when $B = M \times G$], then it can be written as

$$\mathcal{O}(x) = \mathcal{O}[g(x), R(x), \dots, \nabla^{p-2} R(x), F(x), \dots, \mathcal{D}^{p-2} F(x)]\,. \tag{98}$$

In particular, there cannot be any dependence upon the background fields Ψ.

Proof: The proof follows [76], with a slight generalization due to the presence of gauge fields that were not considered in that reference. We first consider the case $B = M \times G$. Then our covariance condition implies the conditions

$$\pounds_\xi \mathcal{O} = \sum_{k=0}^{p} \frac{\partial \mathcal{O}}{\partial (D^k g)} D^k \pounds_\xi g + \sum_{k=0}^{p} \frac{\partial \mathcal{O}}{\partial (D^k A)} D^k \pounds_\xi A \tag{99}$$

for any vector field ξ on M, and

$$0 = \sum_{k=0}^{p} \frac{\partial \mathcal{O}}{\partial (D^k A)} D^k \mathcal{D} h \tag{100}$$

for any Lie-algebra valued function h on M. We first analyze the first of these conditions, following [76]. First, we rewrite all D-derivatives of A in terms of ∇-derivatives (where ∇ is the Levi-Civita connection of g), plus additional terms involving D-derivatives of g. Thus, we write

$$\mathcal{O}(x) = \mathcal{O}[g(x), \dots, D^p g(x), A(x), \dots, \nabla^p A(x), \Psi(x)]\,. \tag{101}$$

Next, we eliminate $D^k g$ in favor of C and its D-derivatives, where C is the tensor field defined by

$$C^\mu{}_{\nu\sigma} = -\frac{1}{2} g^{\mu\alpha} (D_\alpha g_{\nu\sigma} - 2 D_{(\nu} g_{\sigma)\alpha})\,. \tag{102}$$

We thereby obtain

$$\mathcal{O}(x) = \mathcal{O}[g(x), C(x), \dots, D^{p-1} C(x), A(x), \dots, \nabla^p A(x), \Psi(x)]\,. \tag{103}$$

Next, we observe that the symmetrized derivatives of C can be rewritten as

$$
\begin{aligned}
D_{(\alpha_1} \cdots D_{\alpha_l)} C^{\mu}{}_{\gamma\delta} &= D_{(\alpha_1} \cdots D_{\alpha_l} C^{\mu}{}_{\gamma\delta)} \\
&+ \frac{l+3}{4(l+1)(l+2)} \sum_i \nabla_{(\alpha_1} \cdots \widehat{\nabla}_{\alpha_i} \cdots \nabla_{\alpha_l} (R^{\mu}{}_{\gamma\alpha_i\delta} + R^{\mu}{}_{\delta\alpha_i\gamma}) \\
&+ \frac{3l+4}{8(l+1)(l+2)} \sum_{i \neq j} \Big(\nabla_{(\gamma} \nabla_{\alpha_1} \cdots \widehat{\nabla}_{\alpha_i} \widehat{\nabla}_{\alpha_j} \cdots \nabla_{\alpha_l)} R^{\mu}{}_{\alpha_i\delta\alpha_j} \\
&+ \nabla_{(\delta} \nabla_{\alpha_1} \cdots \widehat{\nabla}_{\alpha_i} \widehat{\nabla}_{\alpha_j} \cdots \nabla_{\alpha_l)} R^{\mu}{}_{\alpha_i\gamma\alpha_j} \Big) \\
&+ \text{terms with less than } l \text{ derivatives on } C.
\end{aligned} \tag{104}
$$

By iterating this substitutions, we can achieve that all derivatives of $C^{\mu}{}_{\nu\sigma}$ in $\mathcal{O}$ only appear in totally symmetrized form $D_{(\alpha_1} \dots D_{\alpha_l} C^{\mu}{}_{\nu\sigma)}$, at the expense of possibly having an additional dependence upon the curvature tensor $R^{\mu}{}_{\alpha\beta\gamma}$ of the metric and its covariant derivatives. In other words, we may assume that $\mathcal{O}$ is given as

$$
\begin{aligned}
\mathcal{O} &= \mathcal{O}\Big[g_{\mu\nu}, C^{\mu}{}_{\nu\sigma}, \dots, D_{(\alpha_1} \dots D_{\alpha_{p-1}} C^{\mu}{}_{\nu\sigma)}, \\
&\quad R^{\mu}{}_{\nu\sigma\rho}, \dots, \nabla_{(\alpha_1} \dots \nabla_{\alpha_{p-2})} R^{\mu}{}_{\nu\sigma\rho}, A_{\mu}, \dots, \nabla_{(\alpha_1} \dots \nabla_{\alpha_p)} A_{\mu}; \Phi \Big].
\end{aligned} \tag{105}
$$

We now apply the condition (99) to this expression. We find

$$
\begin{aligned}
& \sum_{k=0}^{p-1} \frac{\partial \mathcal{O}}{\partial (D_{(\alpha_1} \dots D_{\alpha_k} C^{\mu}{}_{\nu\sigma)})} \pounds_{\xi} D_{(\alpha_1} \dots D_{\alpha_k} C^{\mu}{}_{\nu\sigma)} + \frac{\partial \mathcal{O}}{\partial \Psi} \pounds_{\xi} \Psi \\
= & \sum_{k=0}^{p-1} \frac{\partial \mathcal{O}}{\partial (D_{(\alpha_1} \dots D_{\alpha_k} C^{\mu}{}_{\nu\sigma)})} D_{(\alpha_1} \dots D_{\alpha_k} \delta C^{\mu}{}_{\nu\sigma)}
\end{aligned} \tag{106}
$$

where $\delta C^{\mu}{}_{\nu\sigma}$ is the variation arising from the variation $\delta g_{\mu} = \pounds_{\xi} g_{\mu\nu} = 2\nabla_{(\mu}\xi_{\nu)}$ under an infinitesimal diffeomorphism,

$$
\delta C^{\alpha}{}_{\beta\gamma} = g^{\alpha\delta} (D_{(\beta} D_{\gamma)} \xi_{\delta} - \mathcal{R}_{\delta(\beta\gamma)\rho} \xi^{\rho}) - 2D^{(\alpha} \xi^{\delta)} g_{\delta\rho} C^{\rho}{}_{\beta\gamma}, \tag{107}
$$

with $\mathcal{R}_{\mu\nu\sigma\rho}$ the curvature of D. The terms in the above equation arising from an infinitesimal variation of $g_{\mu\nu}, A_{\mu}, R_{\mu\nu\sigma\rho}$ and their ∇-derivatives cancel out. The key point about the above equation is now that, on the left side, there appears no more than one derivative of ξ^{μ}, while on the right side there can appear as many as $p+1$ symmetrized derivatives of ξ^{μ}. Since the symmetrized derivatives of ξ^{μ} can be chosen independently at each given point x in M, it follows that a necessary condition for eq. (112) to hold is that

$$
\frac{\partial \mathcal{O}}{\partial (D_{(\alpha_1} \dots D_{\alpha_k} C^{\mu}{}_{\nu\sigma)})} = 0 \tag{108}
$$

for $k=0,\dots,p-1$. Thus, our expression for $\mathcal{O}$ must have the form

$$\mathcal{O}=\mathcal{O}\left[g_{\mu\nu},R^{\mu}{}_{\nu\sigma\rho},\dots,\nabla_{(\alpha_1}\dots\nabla_{\alpha_{p-2})}R^{\mu}{}_{\nu\sigma\rho},A_{\mu},\dots,\nabla_{(\alpha_1}\dots\nabla_{\alpha_p)}A_{\mu};\Psi\right]. \tag{109}$$

We also get the condition that $\partial\mathcal{O}/\partial\Psi\cdot £_{\xi}\Psi=0$. If Ψ only consists of scalar fields, then it follows immediately that $\mathcal{O}$ cannot have any dependence on Ψ. If Ψ contains tensor fields, then we may reduce this to the situation of only scalar fields by picking a coordinate system, and by treating the coordinate components of Ψ as scalars.

We finally use the condition to show that the A-dependence of $\mathcal{O}$ can only be through the field strength tensor and its covariant derivatives. To show this, we rewrite

$$\begin{aligned}\nabla_{(\alpha_1}\dots\nabla_{\alpha_l)}A_{\mu} &= \nabla_{(\alpha_1}\dots\nabla_{\alpha_l}A_{\mu)}+\frac{l}{l+1}\mathcal{D}_{(\alpha_1}\dots\mathcal{D}_{\alpha_{l-1}}F_{\alpha_l)\mu}\\ &+ \text{terms with no more than } l-1 \text{ derivatives of } A.\end{aligned} \tag{110}$$

By repeatedly substituting this relation into $\mathcal{O}$, we can rewrite it as

$$\begin{aligned}\mathcal{O} &= \mathcal{O}\Big[g_{\mu\nu},F_{\mu\nu},\dots,\mathcal{D}_{(\alpha_1}\dots\mathcal{D}_{\alpha_{p-1})}F_{\mu\nu},\\ &\quad R^{\mu}{}_{\nu\sigma\rho},\dots,\nabla_{(\alpha_1}\dots\nabla_{\alpha_{p-2})}R^{\mu}{}_{\nu\sigma\rho},A_{\mu},\dots,\nabla_{(\alpha_1}\dots\nabla_{\alpha_p}A_{\mu)}\Big].\end{aligned} \tag{111}$$

We now substitute this into the (infinitesimal version) of our condition (2.3), to get

$$\begin{aligned}0 &= \sum_{k=0}^{p}\frac{\partial\mathcal{O}}{\partial(\nabla_{(\alpha_1}\dots\nabla_{\alpha_k}A_{\mu)})}\nabla_{(\alpha_1}\dots\nabla_{\alpha_k}\mathcal{D}_{\mu)}h\\ &+ \sum_{k=0}^{p-2}\frac{\partial\mathcal{O}}{\partial(\mathcal{D}_{(\alpha_1}\dots\mathcal{D}_{\alpha_k)}F_{\mu\nu})}[h,\mathcal{D}_{(\alpha_1}\dots\mathcal{D}_{\alpha_k)}F_{\mu\nu}],\end{aligned} \tag{112}$$

for all Lie-algebra valued functions h. Note that, in the second sum, we have no derivatives of h, while in the first sum we have at least one symmetrized derivative of h. Since the symmetrized derivatives of h are independent at each point, the above equation can only hold if

$$\frac{\partial\mathcal{O}}{\partial(\nabla_{(\alpha_1}\dots\nabla_{\alpha_k}A_{\mu)})}=0 \tag{113}$$

for all k. This proves the Thomas replacement theorem in the case when $B=M\times G$. But, since it is a local statement and any principal fibre bundle is locally trivial, it must in fact hold for any principal fibre bundle. $\square$

3 Quantized field theories on curved spacetime: Renormalization

3.1 Definition of the free field algebra $\mathbf{W}_0$ for scalar field theory

Consider a classical scalar field ϕ described by the quadratic Lagrangian

$$\mathbf{L}_0 = \frac{1}{2}(d\phi \wedge *d\phi - m^2 * \phi^2) \, . \tag{114}$$

The quantity m^2 is a real parameter (we do not assume $m^2 \geq 0$). In this section, we explain how to quantize such a theory in curved spacetime, and how to define Wick powers and time-ordered products of ϕ at the quantum level. We assume only that (M,g) is globally hyperbolic and we assume for the rest of the paper that the spacetime dimension is 4. We do not assume that (M,g) has any symmetries. As discussed above, if (M,g) is globally hyperbolic, then the Klein-Gordon equation has a well-posed initial value formulation and unique retarded and advanced propagators Δ_R and Δ_A. A fundamental object in the quantization of ϕ is the commutator function,

$$\Delta = \Delta_A - \Delta_R \tag{115}$$

which is antisymmetric, $\Delta(x,y) = -\Delta(y,x)$. We want to define a non-commutative product $\star_\hbar$ between classical field observables such that

$$\phi(x) \star_\hbar \phi(y) - \phi(y) \star_\hbar \phi(x) = i\hbar\Delta(x,y)\mathbb{1} \, . \tag{116}$$

This formula is motivated by the fact that, as $\hbar \to 0$, we would like the above commutator divided by $i\hbar$ to go to the classical Peierls bracket. The classical Peierls bracket for a linear scalar field with Lagrangian $\mathbf{L}_0$, however, is given by $\{\phi(x), \phi(y)\}_{\text{P.B.}} = \Delta(x,y)$, see e.g. [39].

To define the desired "deformation quantization", we proceed as follows. We first consider the free *-algebra generated by the expressions $\phi(f)$, where f is any smooth compactly supported testfunction, to be thought of informally as the integral expressions $\int \phi(x) f(x) \, dx$. We now simply factor this free algebra by the relation (116). This defines the desired deformation quantization algebra $\mathbf{W}_{00}$. Evidently, the construction of $\mathbf{W}_{00}$ only depends upon the spacetime (M,g) and its orientations, because these data uniquely determine the retarded and advanced propagators.

The algebra $\mathbf{W}_{00}$ by itself is too small to serve as an arena for renormalized perturbation theory. It does not, for example, even contain the Wick-powers of the free field, or other quantized composite fields, which are a minimal input to even define interactions at the quantum level. More generally, to do perturbation theory we need an algebra that also contains the time-ordered products of composite fields, and these are, of course, not contained in $\mathbf{W}_{00}$ either. Thus, our first task is to define an algebra that is sufficiently big to contain such quantities. The key input in the construction of such an algebra is an arbitrary, but fixed 2-point function $\omega(x,y)$

on $M \times M$ of "Hadamard type" which serves to define a suitable completion of $\mathbf{W}_{00}$. This is by definition a distribution on $M \times M$ which is (a) a bisolution to the equations of motion, that is,

$$(\Box - m^2)_x \omega(x,y) = (\Box - m^2)_y \omega(x,y) = 0\,, \tag{117}$$

which (b) satisfies

$$\omega(x,y) - \omega(y,x) = i\Delta(x,y) \tag{118}$$

and which (c) has a wave front set [72] of "Hadamard type" [93]

$$\begin{aligned} \mathrm{WF}(\omega) \;\; = \;\; & \{(x_1,k_1,x_2,k_2) \in T^*M \times T^*M; \\ & x_1 \text{ and } x_2 \text{ can be joined by null-geodesic } \gamma \\ & k_1 = \dot\gamma(0) \text{ and } k_2 = -\dot\gamma(1), \text{ and } k_1 \in \bar{V}^+\}\,. \end{aligned} \tag{119}$$

The wave front set completely characterizes the singularity structure of ω, and its definition and properties are recalled in appendix C. It can be shown that, on any globally hyperbolic spacetime (M,g), there exist infinitely many distributions ω of Hadamard type [78, 50, 81]. Using ω, we now define the following set of generators of $\mathbf{W}_{00}$, where $u = f_1 \otimes \cdots \otimes f_n$:

$$\begin{aligned} F(u) \;\; &= \;\; \int_M \cdots \int_M f_1(x_1)\cdots f_n(x_n) :\phi(x_1)\cdots\phi(x_n):_\omega \, dx_1 \dots dx_n \\ &= \;\; \frac{d^n}{d\tau_1 \dots d\tau_n} \exp_{\star_\hbar}\Big(i\sum_j \tau_j \phi(f_j) + \frac{\hbar}{2}\sum_{i,j} \tau_i\tau_j \omega(f_i,f_j)\Big)\Big|_{s_i=0}\,. \end{aligned} \tag{120}$$

The commutator property of ω implies that the quantities $:\phi(x_1)\dots\phi(x_n):_\omega$ are symmetric in its arguments. In fact, these quantities are nothing but the "normal ordered field products" (with respect to ω), but we note that we do not think of these objects as operators defined on a Hilbert space as is usually done when introducing normal ordered expressions.

So far, we have done nothing but to introduce a new set of expressions in $\mathbf{W}_{00}$ that generate this algebra. We can express the product between to elements $F(u), F(v)$ of the form (120) as

$$F(u) \star_\hbar F(v) = \sum_k \hbar^k F(u \otimes_k v) \tag{121}$$

where $u \otimes_k v$ is the k-times contracted tensor product of distributions u, v in n resp. m spacetime variables. It is defined by

$$\begin{aligned} &(u \otimes_k v)(x_1, \dots, x_{n+m-2k}) = \\ &\quad \frac{n!m!}{k!}\sum_\pi \int u(x_{\pi(1)}, \dots, y_1, \dots) v(x_{\pi(n-k+1)}, \dots, y_{k+1}, \dots) \prod_{i=1}^k \omega(y_i, y_{k+i}) dy_1 \dots dy_{2k}\,, \end{aligned} \tag{122}$$

where the sum is over all permutations of $n+m-2k$ elements. A somewhat more symbolic, but more compact and suggestive way to write the product is

$$F(u) \star_\hbar F(v) =: F(u) \exp\Big(\frac{1}{2}\hbar_{<\mathcal{D}>}\Big) F(v) :_\omega \tag{123}$$

where $_<\mathcal{D}_>$ is the bi-differential operator defined by

$$_<\mathcal{D}_> = \int \frac{\delta_L}{\delta\phi(x)} \omega(x,y) \frac{\delta_R}{\delta\phi(y)} \, dxdy \, . \tag{124}$$

The superscripts on the functional derivatives indicate that the first derivative acts to the left, and the second one to the right factor in a tensor product. These functional derivatives are to be understood to act on an expression like $:\phi(x_1)\dots\phi(x_n):_\omega$ a classical product of classical fields in $\mathbf{P}(M)$. The point is now that the product can still be defined on a much larger class of expressions. These expressions are of the form

$$F(u) = \int u(x_1,\dots,x_n) :\phi(x_1)\cdots\phi(x_n):_\omega \, dx_1 \dots dx_n \quad (n \geq 1) \, , \tag{125}$$

where u is now a *distribution* on M^n, rather than the product of n smooth functions on M as above in eq. (120). To make the product well defined, we only need to impose a mild wave-front set condition on the u [39]:

$$\mathrm{WF}(u) \cap \bigcup_{x \in M} [(\bar{V}_x^+)^{\times n} \cup (\bar{V}_x^-)^{\times n}] = \emptyset \, , \tag{126}$$

with $\bar{V}_x^\pm$ denoting the closure of the future/past lightcone at x. The reason for imposing this condition is that it ensures, together with (119), that the distributional products in the contracted tensor products that arise when carrying out the product $F \star_\hbar G$ of two expressions of the type (125) make sense. The point is that in such a product, there appear distributional products of u, v, ω in the contracted tensor product of u, v, see eq. (122). Normally, the product of distributions does not make sense, but due to our wave front set conditions on u, v, ω, the relevant products exist due to the fact that vectors in the wave front set of ω, u, v can never add up to 0, see appendix C for details. We define the desired enlarged algebra, $\mathbf{W}_0$, to be the algebra generated by (125), with the product $\star_\hbar$. It can be viewed in a certain sense as the closure of $\mathbf{W}_{00}$, because the distributions u in eq. (125) can be approximated, to arbitrarily good precision by sums of smooth functions of the form $f_1 \otimes \cdots \otimes f_n$ as in (120) (in the Hormander topology [72]). The algebra $\mathbf{W}_0$ will turn out to be big enough to serve as an arena for perturbation theory. For example, it can be seen immediately that $\mathbf{W}_0$ contains normal ordered Wick-powers of $\phi(x)$: Namely, since the wave-front set of the delta-distribution on M^n is

$$\mathrm{WF}(\delta) = \{(x,k_1,\dots,x,k_n); \quad x \in M, k_i \in T_x^*M, \sum k_i = 0\} \tag{127}$$

it follows that $u(y,x_1,\dots,x_n) = f(y)\delta(y,x_1,\dots,x_n)$ satisfies the wave front condition (126). The corresponding generator F as in (125) may be viewed as the normal ordered Wick power $:\phi^n(x):_\omega$, smeared with $f(x)$.

As it stands, the Klein-Gordon equation is not implemented in the algebra $(\mathbf{W}_0, \star_\hbar)$. This could easily be incorporated by factoring $\mathbf{W}_0$ by an appropriate ideal (i.e., a linear subspace that

is stable under $\star_\hbar$-multiplication by any $F \in \mathbf{W}_0$). The ideal for the field equation is simply the linear space

$$\mathcal{J}_0 = \left\{ F = \int u(x_1,\dots,x_n) : \phi(x_1)\cdots \frac{\delta S_0}{\delta\phi(x_i)} \cdots \phi(x_n) :_\omega dx_1 \dots dx_n, \right.$$
$$\left. \text{for some } u \text{ of compact support, } \mathrm{WF}(u) \cap \bigcup_{x\in M} [(\bar{V}_x^+)^{\times n} \cup (\bar{V}_x^-)^{\times n}] = \emptyset \right\} \quad (128)$$

of generators containing a factor of the wave equation. This space is stable under the adjoint operation and $\star_\hbar$-products with any $F \in \mathbf{W}_0$ by eq. (117) and so indeed an ideal. If we consider the factor algebra

$$pr : \mathbf{W}_0 \to \mathcal{F}_0 = \mathbf{W}_0 / \mathcal{J}_0 \,, \quad (129)$$

then within $\mathcal{F}_0$, the field equation $(\Box - m^2)\phi(x) = 0$ holds. The factor algebra $\mathcal{F}_0$ is the algebra of physical interest for free field theory. For physical applications, one is interested in representations of $\mathcal{F}_0$ as operators on a Hilbert space, $\mathcal{H}_0$, and in n-point functions of observables in $\mathcal{F}_0$ in physical states. However, in the context of perturbation theory, it will be much more useful to work with the algebra $\mathbf{W}_0$ at intermediate stages.

To make physical predictions, one finally needs to represent the algebra of observables $\mathcal{F}_0$ as linear operators with a dense, invariant domain on a Hilbert space $\mathcal{H}_0$. A vector state $|\Psi\rangle$ in $\mathcal{H}_0$ is said to be of Hadamard form if its n-point functions

$$G_n^\Psi(x_1,\dots,x_n) = \langle \Psi | \pi_0(\phi(x_)) \dots \pi_0(\phi(x_n)) | \Psi \rangle \quad (130)$$

are of "Hadamard form". By this one means that the 2-point function has a wave front set of Hadamard form (119), and that its truncated n-point functions[4] are smooth for $n \neq 2$. A Hadamard representation is a representation containing a dense, invariant domain of Hadamard states. Hadamard representations may be constructed on any globally hyperbolic spacetime as one may show using the deformation argument of [49, 81] (or the construction of [78], and combining these with those of [65]). We describe the deformation construction below in sec. 4.2 in the context of gauge theories.

It is clear that, since $\mathbf{W}_0(M,g)$ was obtained as the completion of the algebra $\mathbf{W}_{00}(M,g)$, also $\mathbf{W}_0(M,g)$ depends locally and covariantly upon the metric. Because this fact will be of key importance when we formulate the local and covariance condition of renormalized time-ordered products, we now explain more formally what exactly we mean by this statement. Consider two oriented and time-oriented spacetimes (M,g) and (M',g') and a map $\psi : M \to M'$ which is an orientation and causality preserving[5] isometric embedding. Then there is a corresponding

[4]The truncated n-point functions of a hierarchy of n-point distributions $\{h_n\}$ are defined by the generating functional $h^c(\mathrm{e}_\otimes^f) = \log h(\mathrm{e}_\otimes^f)$, where $h(\mathrm{e}_\otimes^f) = \sum_n h_n(f,f,\dots,f)/n!$.

[5]An isometric embedding may be such that the intrinsic notion of causality is not the same as the notion of causality inherited from the ambient space. Examples of this sort may be constructed by embedding suitable regions of Minkowski spacetime into Minkowski space with periodic identifications in one or more spatial directions.

isomorphism

$$\alpha_\psi : \mathbf{W}_0(M,g) \to \mathbf{W}_0(M',g')\,, \tag{131}$$

which behaves naturally under composition of embeddings. This map is simply defined on $\mathbf{W}_{00}(M,g)$ by setting $\alpha_\psi(\phi_{M,g}(f)) = \phi_{M',g'}(\psi_* f)$, where $\psi_* f(x') = f(x)$ for $x = \psi(x')$. Since, as explained above, $\mathbf{W}_0(M,g)$ is essentially the closure of $\mathbf{W}_{00}(M,g)$, we can define α_ψ on $\mathbf{W}_0(M,g)$ by continuity. The action of α_ψ on F of the form (120) may be calculated straightforwardly from the definition. However, we note that its form will depend on the choices ω and ω' for the Hadamard bidistributions on M respectively M', and will look somewhat involved if ω and ω' are such that $\psi^*\omega' \neq \omega$. These expressions are given in [62], but will not be needed here.

3.2 Renormalized Wick products and their time-ordered products

In the previous section we have laid the groundwork for the construction of linear quantum field theory in curved spacetime by giving the definition of an algebra $\mathbf{W}_0(M,g)$ associated with a free Lagrangian $\mathbf{L}_0$ that can be viewed as a deformation quantization of the algebra of classical observables with the Peierls bracket. In this section we shall identify, within $\mathbf{W}_0(M,g)$, the various objects that have the interpretation of the various Wick powers in the theory, and their time-ordered products. Those objects will be the quantities of prime interest in the perturbative constructions in the subsequent sections. For simplicity, we first address the case when $\mathbf{L}_0$ describes a linear, hermitian scalar field ϕ, see eq. (114).

Actually, for reasons that we will explain below, it is convenient to adopt a unified viewpoint on the Wick products and their time-ordered products. We define a time-ordered product with n factors (where $n \geq 1$) to be a linear map

$$T_n : \mathbf{P}^{k_1}(M) \otimes \dots \mathbf{P}^{k_n}(M) \to \mathcal{D}'\Big(M^n; \wedge^{k_1} T^*M \times \dots \times \wedge^{k_n} T^*M\Big) \otimes \mathbf{W}_0\,, \tag{132}$$

taking values in the distributions over M^n with target space $\mathbf{W}_0$. Thus, the linear map T_n takes as arguments the tensor product of n local covariant classical forms $\mathcal{O}_1, \dots, \mathcal{O}_n$, and it gives an expression $T_n(\mathcal{O}_1(x_1) \otimes \dots \otimes \mathcal{O}_n(x_n))$, which is itself a distribution in n spacetime variables $x_1, \dots, x_n$, with values in $\mathbf{W}_0$, i.e., $T_n(\mathcal{O}_1(x_1) \otimes \dots \otimes \mathcal{O}_n(x_n))$ is itself a map that needs to be smeared with n-test forms $f_1(x_1), \dots, f_n(x_n)$, where the i-th test form is an element in the set of compactly supported smooth forms $f_i \in \Omega_0^{4-k_i}(M)$ over M. The set $\mathcal{D}'(M^n; \wedge^{k_1} T^*M \times \dots \times \wedge^{k_n} T^*M)$ denotes the dual space (in the standard distribution topology [72]) of the space of forms $\Omega_0^{4-k_1}(M) \times \dots \times \Omega_0^{4-k_n}(M)$.

The time-ordered products T_n are characterized abstractly by certain properties which we will list. We define the Wick powers of a field to be the time-ordered products with 1 factor, i.e., $n = 1$. We will formulate the properties of the time-ordered products in the form of axioms in this section, but we will see in the following section that one can turn these properties into a concrete constructive algorithm for these quantities. In fact, as we will see, the properties

that we wish the time-ordered products to have do not *uniquely* characterize them, but leave a certain ambiguity. This ambiguity corresponds precisely to the renormalization ambiguity in other approaches in flat spacetime, with the addition of couplings to curvature. However, we note that our time-ordered products are rigorously defined, by contrast to the corresponding quantities in other approaches to renormalization in flat spacetime, where they are a priori only formal (i.e., infinite) objects.

T1 Locality and covariance The time ordered products are locally and covariantly constructed in terms of the metric. This means that, if $\psi : M \to M'$ is a causality preserving isometric embedding between two spacetimes preserving the causal structure, and α_ψ denotes the corresponding homomorphism $\mathbf{W}_0(M,g) \to \mathbf{W}_0(M',g')$, see eq. (131), then we have

$$\alpha_\psi \circ T_n = T'_n \circ \bigotimes^n \psi_* \tag{133}$$

where T_n denotes the time-ordered product on (M,g), while T'_n denotes the time-ordered product on (M',g'). The mapping $\psi_* : \mathbf{P}(M) \to \mathbf{P}(M')$ is the natural push-forward map. Thus, the local and covariance condition imposes a relation between the construction of time-ordered products on locally isometric spacetimes.Written more explicitly (in the case of scalar operators), the local covariance condition is

$$\alpha_\psi \Big[T_n(\phi^{k_1}(x_1) \otimes \dots \phi^{k_n}(x_n)) \Big] = T'_n(\phi^{k_1}(x'_1) \otimes \dots \phi^{k_n}(x'_n)) \quad \psi(x_i) = x'_i \,. \tag{134}$$

In particular, if $n = 1$, then the Wick products $T_1(\mathcal{O}(x))$ are local covariant fields in one variable. As we will see more clearly in the next subsection, the requirement of locality and covariance is a non-trivial renormalization condition already in the case of 1 factor.

It is instructive to consider the local covariance requirement for the special case where $M = M'$ is Minkowksi spacetime, with $g = g'$ the Minkowski metric $-dt^2 + dx^2 + dy^2 + dz^2$. In that case, the causality and orientation preserving isometric embeddings are just the proper, orthochronous Poincare transformations $\psi = (\Lambda, a) \in P_+^\uparrow$, while the map α_ψ may be implemented by $Ad(U_0(\Lambda,a))$ in the vacuum Hilbert space representation π_0 of the algebra $\mathbf{W}_0$ (we need to assume $m^2 \geq 0$ to have that representation), with $U_0(\Lambda,a)$ the unitary representative of the proper orthochronous Poincare transformation (Λ,a) on the Hilbert space of the representation π_0. The local covariance condition (135) reduces in that case to

$$Ad[U_0(\Lambda,a)]\pi_0\Big(T_n\big(\phi^{k_1}(x_1) \otimes \dots \phi^{k_n}(x_n)\big)\Big) = \pi_0\Big(T_n\big(\phi^{k_1}(\Lambda x_1 - a) \otimes \dots \phi^{k_n}(\Lambda x_n - a)\big)\Big) \tag{135}$$

which is the standard transformation law for the time ordered product (and in fact any relativistic field) in Minkowski spacetime.

T2 Scaling. We would like the time-ordered products to satisfy a certain scaling relation. For distributions $u(x), x \in \mathbb{R}^n$ on flat space, it is natural to consider the scaled distribution $u(\mu x), \mu \in \mathbb{R}_+$. Such a distribution is then said to scale homogeneously with degree D if $u(\mu x) = \mu^D u(x)$, in the sense of distributions, which is equivalent to the differential relation

$$\Big(\mu\frac{\partial}{\partial\mu} - D\Big) u(\mu x) = 0\,. \tag{136}$$

More generally, it is said to scale "polyhomogeneously" or "homogeneously up to logarithms" if instead only

$$\Big(\mu\frac{\partial}{\partial\mu} - D\Big)^N u(\mu x) = \frac{\partial^N}{\partial(\log\mu)^N}\left[\mu^D u(\mu x)\right] = 0\,. \tag{137}$$

holds for some $N \geq 2$, which gives the highest power $+1$ of the logarithmic corrections.

For the quantities in the quantum field theory associated with the Lagrangian $\mathbf{L}_0$ on a generic curved spacetime without dilation symmetry, we do not expect a simple scaling behavior under rescalings in an arbitrarily chosen coordinate system. However, we know that the Lagrangian $\mathbf{L}_0$ has an invariance under a rescaling

$$g \mapsto \mu^2 g, \quad m^2 \to \mu^{-2} m^2, \quad \phi \mapsto \mu^{-1}\phi\,. \tag{138}$$

It is therefore natural to expect that the time-ordered products can be constructed so as to have a simple scaling behavior under such a rescaling. However, due to quantum effects, one cannot expect an exactly homogeneous scaling, but only a homogeneous scaling behavior that is modified by logarithms. To describe this behavior, we must first take into account that the time-ordered products associated with the spacetime metric g live in a different algebra than the time-ordered products associated with $\mu^2 g$, so we must first identify these algebras. This is achieved by the linear map $\sigma_\mu : \phi \mapsto \mu\phi$, which may be checked to define an isomorphism between $\mathbf{W}_0(M, g, m^2)$ and $\mathbf{W}_0(M, \mu^2 g, \mu^{-2} m^2)$. The desired polyhomogeneous scaling behavior is then formulated as follows. Let

$$T_n[\mu] = \sigma_\mu^{-1} \circ T_n \circ \bigotimes^n \exp(\ln\mu \cdot \mathcal{N}_d) \tag{139}$$

where $\mathcal{N}_d$ is the dimension counter, defined as $\mathcal{N}_d := \mathcal{N}_c + \mathcal{N}_f + \mathcal{N}_r$, where $\mathcal{N}_c, \mathcal{N}_f, \mathcal{N}_r : \mathbf{P}(M) \to \mathbf{P}(M)$ are the number counting operators for the coupling constants, fields, and curvature terms, defined for Klein-Gordon theory in 4 spacetime dimensions by

$$\mathcal{N}_f \;:=\; \sum_k (1+k)(\nabla^k\phi)\frac{\partial}{\partial(\nabla^k\phi)}\,, \tag{140}$$

$$\mathcal{N}_c \;:=\; 2m^2\frac{\partial}{\partial m^2}\,, \tag{141}$$

$$\mathcal{N}_r \;:=\; \sum_k (k+2)(\nabla^k R)\frac{\partial}{\partial(\nabla^k R)}\,. \tag{142}$$

For example

$$T_n[\mu](\phi^{k_1}(x_1)\otimes\cdots\otimes\phi^{k_n}(x_n)) = \mu^{k_1+\cdots+k_n}\sigma_\mu^{-1} T_n(\phi^{k_1}(x_1)\otimes\cdots\otimes\phi^{k_n}(x_n)). \tag{143}$$

Because we have put the identification map r_μ on the right side, $T_n[\mu]$ defines a new time ordered product in the algebra associated with the unscaled metric, g, and coupling constants. In the absence of scaling anomalies, this would be equal to the original T_n for all $\mu \in \mathbb{R}_+$. As we have said, it is not possible to achieve this exactly homogeneous scaling behavior, so we only postulate the polyhomogeneous scaling behavior

$$\frac{\partial^N}{\partial(\log\mu)^N} T_n[\mu] = 0. \tag{144}$$

T3 Microlocal Spectrum condition. Consider a time ordered product $T_n(\mathcal{O}_1(x_1)\otimes\cdots\otimes\mathcal{O}_n(x_n))$ as an $\mathbf{W}_0$ valued distribution on M^n. Then we require that

$$\mathrm{WF}(T_n) \subset C_T(M,g), \tag{145}$$

where the set $C_T(M,g) \subset T^*M^n \setminus 0$ is described as follows (we use the graph theoretical notation introduced in [17, 18]): Let $G(p)$ be a "decorated embedded graph" in (M,g). By this we mean an embedded graph $\subset M$ whose vertices are points $x_1,\dots,x_n \in M$ and whose edges, e, are oriented null-geodesic curves. Each such null geodesic is equipped with a coparallel, cotangent covectorfield p_e. If e is an edge in $G(p)$ connecting the points x_i and x_j with $i<j$, then $s(e)=i$ is its source and $t(e)=j$ its target. It is required that p_e is future/past directed if $x_{s(e)} \notin J^\pm(x_{t(e)})$. With this notation, we define

$$\begin{aligned} C_T(M,g) \;\; = \;\; & \Big\{(x_1,k_1;\dots;x_n,k_n) \in T^*M^n \setminus 0 \mid \exists \text{ decorated graph } G(p) \text{ with vertices} \\ & x_1,\dots,x_n \text{ such that } k_i = \sum_{e:s(e)=i} p_e - \sum_{e:t(e)=i} p_e \quad \forall i\Big\}. \end{aligned} \tag{146}$$

T4 Smoothness. The functional dependence of the time ordered products on the spacetime metric, g, is such that if the metric is varied smoothly, then the time ordered products vary smoothly, in the sense described in [62].

T5 Analyticity. Similarly, we require that, for an analytic family of analytic metrics (depending analytically upon a set of parameters), the expectation value of the time-ordered products in an analytic family of states[6] varies analytically in the same sense as in T4.

[6]As explained in remark (2) on P. 311 of [62], it suffices to consider a suitable analytic family of linear functionals on $\mathbf{W}_0$ that do not necessarily satisfy the positivity condition required for states.

T6 Symmetry. The time ordered products are symmetric under a permutation of the factors,

$$T_n(\mathcal{O}_1(x_1)\otimes\cdots\otimes\mathcal{O}_n(x_n)) = T_n(\mathcal{O}_{\pi 1}(x_{\pi 1})\otimes\cdots\otimes\mathcal{O}_{\pi n}(x_{\pi n})) \tag{147}$$

for any permutation π.

T7 Unitarity. Let $\bar{T}_n(\otimes_i \mathcal{O}_i(x_i)) = [T_n(\otimes_i \mathcal{O}_i(x_i)^*)]^*$ be the "anti-time-ordered" product. Then we require

$$\bar{T}_n\Big(\bigotimes_{i=1}^n \mathcal{O}_i(x_i)\Big) = \sum_{I_1\sqcup\cdots\sqcup I_j=\underline{n}} (-1)^{n+j} T_{|I_1|}\Big(\bigotimes_{i\in I_1}\mathcal{O}_i(x_i)\Big)\star_\hbar\ldots\star_\hbar T_{|I_j|}\Big(\bigotimes_{j\in I_j}\mathcal{O}_j(x_j)\Big), \tag{148}$$

where the sum runs over all partitions of the set $\{1,\ldots,n\}$ into pairwise disjoint subsets $I_1,\ldots,I_j$.

T8 Causal Factorization. The "product" T_n is time ordered in the sense that the following causal factorization property is to be satisfied. Let $\{x_1,\ldots,x_i\}\cap J^-(\{x_{i+1},\ldots,x_n\})=\emptyset$. Then we have

$$\begin{aligned} &T_n(\mathcal{O}_1(x_1)\otimes\cdots\otimes\mathcal{O}_n(x_n)) \\ &\qquad = T_i(\mathcal{O}_1(x_1)\otimes\cdots\otimes\mathcal{O}_i(x_i))\star_\hbar T_{n-i}(\mathcal{O}_{i+1}(x_{i+1})\otimes\cdots\otimes\mathcal{O}_n(x_n)). \end{aligned} \tag{149}$$

For the case of 2 factors, this means

$$T_2(\mathcal{O}_1(x_1)\otimes\mathcal{O}_2(x_2)) = \begin{cases} T_1(\mathcal{O}_1(x_1))\star_\hbar T_1(\mathcal{O}_2(x_2)) & \text{when } x_1\notin J^-(x_2); \\ T_1(\mathcal{O}_2(x_2))\star_\hbar T_1(\mathcal{O}_1(x_1)) & \text{when } x_2\notin J^-(x_1). \end{cases} \tag{150}$$

T9 Commutator. The commutator of a time-ordered product with a free field is given by lower order time-ordered products times suitable commutator functions, namely

$$\Big[T_n\Big(\bigotimes_i^n \mathcal{O}_i(x_i)\Big),\phi(x)\Big]_{\star_\hbar} = i\hbar\sum_{k=1}^n T_n\Big(\mathcal{O}_1(x_1)\otimes\ldots\int\Delta(x,y)\frac{\delta\mathcal{O}_k(x_k)}{\delta\phi(y)}\otimes\ldots\mathcal{O}_n(x_n)\Big), \tag{151}$$

where Δ is the causal propagator.

T10 Field equation. The free field equation $\delta S_0/\delta\phi$ holds in the sense that

$$T_{n+1}\Big(\frac{\delta S_0}{\delta\phi(x)}\otimes\bigotimes_i^n\mathcal{O}_i(x_i)\Big) = \sum_i T_n\Big(\mathcal{O}_1(x_1)\otimes\cdots\frac{\delta\mathcal{O}_i(x_i)}{\delta\phi(x)}\otimes\cdots\mathcal{O}_n(x_n)\Big) \mod \mathcal{J}_0. \tag{152}$$

T11 Action Ward identity If $d_k = dx_k^\mu \wedge \frac{\partial}{\partial x_k^\mu}$ is the exterior differential acting on the k-th spacetime variable, then we have

$$T_n(\mathcal{O}_1(x_1)\cdots\otimes d_k\mathcal{O}(x_k)\cdots\otimes \mathcal{O}_n(x_n)) = d_k T_n(\mathcal{O}_1(x_1)\otimes\cdots\otimes \mathcal{O}(x_n)) \,. \tag{153}$$

Thus, derivatives can be freely pulled inside the time-ordered products.

Condition T11 can be stated as saying that T_n may alternatively be viewed as a linear map $T_n : A^{\otimes n} \to \mathbf{W}_0$ for each n, where A is the space of all local action functionals, i.e., all expressions of the form $F = \int \mathcal{O} \wedge f$, where $f \in \Omega_0^p(M)$ is any p-form of compact support, and where $\mathcal{O} \in \mathbf{P}^{4-p}$. To explain how this comes about, consider the integrated field polynomial $F = \int f \wedge d\mathcal{O}$. It may equivalently be written as $-\int (df) \wedge \mathcal{O}$, so the time ordered product should give the same result for either choice. T11 means that the time ordered products $\int f(x_i)\, T_n(\cdots\otimes d_i\mathcal{O}(x_i)\otimes\dots)$ and $-\int d_i f(x_i)\, T_n(\cdots\otimes \mathcal{O}(x_i)\otimes\dots)$ are equal, where the exterior derivative $d_i = dx_i^\mu \wedge \partial/\partial x_i^\mu$ acts on the i-th spacetime argument. This means that T_n may be viewed as a functional taking as arguments the integrated functionals (or "actions") in A, because it does not matter how F is represented. This is the origin of the name “action Ward identity” for T11. The action Ward identity also means that we may apply the Leibniz rule for derivative of quantum Wick powers, i.e., time ordered products with one factor, which is why the same condition was called “Leibniz rule” in [66].

3.3 Inductive construction of time-ordered products

In the previous subsection, we have given a list of properties of the local Wick powers and their time-ordered products. We now present an algorithm showing how these can be constructed, and thus in particular demonstrating that axioms T1 through T11 are not empty. We shall reduce the problem to successively simpler problems by a series of reduction steps. These steps are as follows:

1. First, construct the time-ordered products with one factor.

2. Assuming inductively that time-ordered products with n factors have been constructed, we show, following the ideas of “causal perturbation theory” [44, 12, 102, 101] that the time-ordered products with $n+1$ factors are already uniquely fixed, apart from points on the total diagonal, by the lower order time-ordered products.

3. The problem of extending the time-ordered products at order $n+1$ to the total diagonal is reduced to that of extending certain scalar distributions to the total diagonal.

4. The problem of reducing the scalar functions on M^{n+1} to the diagonal is reduced to that of extending a set of distributions on the $(n+1)$-fold Cartesian power of Minkowski space via a curvature expansion.

5. The extension of the Minkowski distributions is performed. This step corresponds to renormalization.

Thus, we shall proceed inductively in the number of factors, n, appearing in the time ordered product $T_n(\mathcal{O}_1(x_1)\otimes\cdots\otimes\mathcal{O}_n(x_n))$. To keep our discussion as simple as possible, we now restrict attention to the case when the fields $\mathcal{O}_i \in \mathbf{P}$ in the time ordered product contain no spacetime derivatives, i.e., $\mathcal{O}_i = \phi^{k_i}$ for some natural numbers k_i. We will also assume for simplicity that external potential v in the Klein-Gordon equation vanishes, so that there are no coupling parameters to consider. We briefly explain how to deal with the general case in the end.

Time-ordered products with 1 factor: For $n = 1$ the time ordered products are just the local covariant Wick powers, i.e., $T_1(\phi^k(x))$ is a local covariant field in one spacetime variable, interpreted as the k-th local covariant Wick power of ϕ. These Wick powers may be constructed as follows. Let $H(x,y)$ be the "local Hadamard parametrix," for the Klein-Gordon operator, given by

$$H(x,y) = \frac{1}{2\pi^2}\left(\frac{u(x,y)}{\sigma + it0} + v(x,y)\,\log(\sigma + it0)\right). \tag{154}$$

Here, $\sigma(x,y)$ is the signed squared geodesic distance between two points x,y in a convex normal neighborhood of M, and u,v are smooth kernels that are locally constructed in terms of the metric, which are determined by the Hadamard recursion relations [24], which are obtained by demanding that H be a bi-solution (modulo a smooth remainder) of the Klein-Gordon equation. Their construction is recalled in Appendix D. The quantity $t(x,y) = T(x) - T(y)$ is defined in terms of an arbitrary global time coordinate T.

Consider now, for any $k \geq 1$, the "locally normal ordered expressions"

$$\begin{aligned} &:\phi(x_1)\cdots\phi(x_k):_{\rm H} \\ &\quad = \frac{\delta^k}{i^k\delta f(x_1)\dots\delta f(x_k)}\exp_{\star_\hbar}\left(i\int_M f(x)\phi(x) + \frac{\hbar}{2}\int_{M\times M} H(x,y)f(x)f(y)\right)\Bigg|_{f=0}. \end{aligned} \tag{155}$$

Because H is defined locally and covariantly in terms of the metric, it follows that $:\phi(x_1)\dots\phi(x_k):_{\rm H}$ are local and covariant fields that are defined in a convex normal neighborhood of the diagonal Δ_k, where

$$\Delta_k = \{(x,x,\dots,x) \mid \quad x \in M\} \subset M^k. \tag{156}$$

The following lemma shows that the normal ordered quantities (155) differ from the quantities $:\phi(x_1)\dots\phi(x_n):_\omega$ only by a smooth function (valued in $\mathbf{W}_0$).

Lemma 5. Let $\omega(x,y)$ be a 2-point function of Hadamard form, i.e., the wave front set $\mathrm{WF}(\omega)$ is given by (119). Then locally (i.e., where H is defined), $\omega - H$ is smooth, i.e.,

$$\omega(x,y) = \frac{1}{2\pi^2}\left(\frac{u(x,y)}{\sigma + it0} + v(x,y)\,\log(\sigma + it0)\right) + \quad \text{(smooth function in } x,y). \tag{157}$$

Furthermore, any two Hadamard states can at most differ by a globally smooth function in x,y.

The proof is given in Appendix E.

Because the normal ordered products may be smeared with a δ-function (or derivatives thereof), we may define

$$T_1\big(\phi^k(x)\big) =: \phi^k(x) :_{\rm H} \tag{158}$$

which is a well defined element in $\mathbf{W}_0$ after smearing with any testfunction $f \in C_0^\infty(M)$. This defines our time-ordered products with one factor. It follows from the definition of H that $T_1(\phi^k(x))$ is a local covariant field, i.e., it satisfies T1 for $n = 1$. The other properties T2—T11 are also seen to be satisfied using the properties of H described in Appendix D.

Time-ordered products with $n > 1$ factors: We have defined the time-ordered products with $n = 1$ factor, and we may inductively assume that time ordered products with properties T1–T11 have been defined for any number of factors $\le n$. The key idea of causal perturbation theory [44, 12, 101, 102] is that the time ordered products with $n+1$ factors are already uniquely determined as algebra-valued distributions on the manifold M^{n+1} minus its total diagonal $\Delta_{n+1} = \{(x,x,\dots,x) \in M^{n+1}\}$ by the causal factorization requirement T8, once the time ordered products with less than or equal to n factors are given. The construction of the time ordered products at order $n+1$ is then equivalent to the task of extending this distribution in a suitable way compatible with the other requirements T1–T10. In order to perform this task in an efficient way, it is useful to derive a number of properties that hold at all orders $m \le n$ as a consequence of T1–T10.

The first property is a local Wick expansion for time ordered products [63]. This is a key simplification, because it will enable one to reduce the problem of extending algebra valued quantities to one of finding an extension of c-number distributions. In the simplest case, when none of the $\mathcal{O}_i$ contain derivatives of ϕ, we have in an open neighborhood of Δ_m

$$\begin{aligned} &T_m\big(\phi^{k_1}(x_1)\otimes\cdots\otimes\phi^{k_m}(x_m)\big) \\ &\qquad = \sum_{0\le j_i\le k_i}\prod_i \binom{k_i}{j_i} t_{j_1,\dots,j_m}(x_1,\dots,x_m) :\phi^{k_1-j_1}(x_1)\cdots\phi^{k_m-j_m}(x_m):_{\rm H} \end{aligned} \tag{159}$$

for all $1 < m \le n$, where $t_{j_1,\dots,j_m}$ are c-number distributions. The Wick expansion when derivatives are present is analogous. The Wick expansion formula can be proved from axiom T9. Because the time-ordered products are local and covariant, the c-number distributions in the Wick expansion have the same property, in the sense that if $\psi : (M',g') \to (M,g)$ is an isometric, causality and orientation preserving embedding, so that if $\psi^* g = g'$, then

$$t_{j_1,\dots,j_m}\big[\psi^* g; x_1,\dots,x_m\big] = t_{j_1,\dots,j_m}\big[g; \psi(x_1),\dots,\psi(x_m)\big]\,. \tag{160}$$

Because H and the local normal ordered products are in general only defined in a neighborhood of the diagonal, it it follows that also the c-number distributions are only defined on a neighborhood of the diagonal, but this will turn out to be sufficient for our purposes.

It follows from the scaling property T2 and the corresponding scaling properties of H that

$$\frac{\partial^N}{\partial(\log\mu)^N}\left\{\mu^{j_1+\cdots+j_m} t_{j_1,\ldots,j_m}\left[\mu^{-2}m^2,\mu^2 g;x_1,\ldots,x_m\right]\right\}=0 \tag{161}$$

for some N. This relation, together with the condition of locality and covariance and the analytic dependence of the time ordered products on the metric, can be used to derive a subsequent "scaling-" or "curvature expansion" [63] of each of the distributions $t_{j_1,\ldots,j_m}$ in powers of the Riemann tensor and the coupling constants (in our case only m^2) at a reference point:

Proposition 0: The distributions $t := t_{j_1,\ldots,j_m}$ have the asymptotic expansion

$$t(\exp_y\xi_1,\ldots,\exp_y\xi_{m-1},y)=\sum_{k=0}^{S} C^k_{\mu_1\ldots\mu_t}(y)\,u_k^{\mu_1\ldots\mu_t}(\xi_1,\ldots,\xi_m)+r^S(y,\xi_1,\ldots,\xi_{m-1})\,. \tag{162}$$

in an open neighborhood of the diagonal Δ_m. The terms have the following properties:

(i) The remainder r^S is a distribution of scaling degree (see Appendix C for the mathematical definition of this concept) strictly lower than the scaling degree of any term in the sum.

(ii) Each u_k is a Lorentz invariant distribution on $(\mathbb{R}^4)^{m-1}$, i.e.,

$$u_k^{\mu_1\ldots\mu_t}(\Lambda\xi_1,\ldots,\Lambda\xi_m)=\Lambda^{\mu_1}_{\nu_1}\cdots\Lambda^{\mu_t}_{\nu_t}u_k^{\nu_1\ldots\nu_t}(\xi_1,\ldots,\xi_m)\quad\forall\Lambda\in SO_0(3,1)\,. \tag{163}$$

(iii) Each distribution u_k scales almost homogeneously under a coordinate rescaling, i.e.,

$$\frac{\partial^N}{\partial(\log\mu)^N}\left[\mu^\rho u_k^{\mu_1\ldots\mu_t}(\mu\xi_1,\ldots,\mu\xi_{m-1})\right]=0 \tag{164}$$

with $\rho\in\mathbb{N}$. The scaling condition can be rewritten equivalently as

$$\Big(\sum_{i=1}^{m-1}\xi_i^\nu\frac{\partial}{\partial\xi_i^\nu}-\rho\Big)^N u_k^{\mu_1\ldots\mu_t}(\xi_1,\ldots,\xi_{m-1})=0\,. \tag{165}$$

(iv) Each term C^k is a polynomial in m^2 and the covariant derivatives of the Riemann tensor ,

$$C^k_{\mu_1\ldots\mu_t}(y)=C^k_{\mu_1\ldots\mu_t}[m^2,R(y),\nabla R(y),\ldots,\nabla^l R(y)]\,. \tag{166}$$

(v) The scaling degree $\rho=sd(u_k)$ is given by

$$sd(u_k)=\sum_i j_i-\mathcal{N}_r(C^k)\,, \tag{167}$$

where $\mathcal{N}_r$ is the dimension counting operator for curvature terms and dimensionful coupling constants (in our case only m^2), see eq. (140).

By the above proposition, we see that, by including sufficiently (but finitely many) terms in the scaling expansion (162) (i.e., choosing S sufficiently large), one can achieve that the remainder r^S has arbitrarily low scaling degree. It does *not* mean that the sum is convergent in any sense (it is not).

Having stated the detailed properties of the time ordered products with $\leq n$ factors, we are now resume the main line of the argument and perform the construction of the time-ordered products with $n+1$ factors. Let I be a proper subset of $\{1,2,\dots,n+1\}$, and let U_I be the subset of M^{n+1} defined by

$$U_I = \{(x_1,x_2,\dots,x_{n+1}) \mid x_i \notin J^-(x_j) \quad \text{for all } i \in I, j \notin I\}. \tag{168}$$

It can be seen [18] that the sets U_I are open and that the collection $\{U_I\}$ of these sets covers the manifold $M^{n+1} \setminus \Delta_{n+1}$. We can therefore define an algebra valued distribution T_{n+1} on this manifold by declaring it for each $(x_1,\dots,x_{n+1}) \in U_I$ by

$$T_{n+1}\big(\phi^{k_1}(x_1) \otimes \cdots \otimes \phi^{k_{n+1}}(x_{n+1})\big) = \\ T_{|I|}\big(\otimes_{i \in I} \phi^{k_i}(x_i)\big) \star_\hbar T_{n+1-|I|}\big(\otimes_{j \in \underline{n+1} \setminus I} \phi^{k_j}(x_j)\big) \quad \forall (x_1,\dots,x_{n+1}) \in U_I. \tag{169}$$

To avoid a potential inconsistency in this definition for points in $U_I \cap U_J \neq \emptyset$ for different I,J, we must show that the definition agrees for different I,J. This can be achieved using the causal factorization property T8 of the time ordered products with less or equal than n factors [44, 18]. Property T8 applied to the time ordered products with $n+1$ factors also implies that the restriction of T_{n+1} to $M^{n+1} \setminus \Delta_{n+1}$ must agree with (169). Thus, property T8 alone determines the time ordered products up to the total diagonal, as we desired to show, see [18] for details.

In fact—assuming that time ordered products with less or equal than n factors have been defined so as to satisfy properties T1–T11 on M^n—one can argue in a relatively straightforwardly way that the fields defined by eq. (169) with $n+1$ factors automatically satisfy[7] the restrictions of properties T1–T9 to $M^{n+1} \setminus \Delta_{n+1}$, while T10 and T11 are empty in the present case for time ordered products without derivatives.

Our remaining task is to find an extension of each of the algebra-valued distributions T_{n+1} in $n+1$ factors from $M^{n+1} \setminus \Delta_{n+1}$ to all of M^{n+1} in such a way that properties T1–T9 continue to hold for the extension. This step, of course, corresponds to renormalization. Condition T8 does not impose any additional conditions on the extension, so we need only satisfy T1–T7 and T9. However, it is not difficult to see that if an extension T_{n+1} is defined that satisfies T1–T5 and T9, then that extension can be modified, if necessary, so as to also satisfy the symmetry and unitarity conditions, T6 and T7, see [62].

Thus, we have reduced the problem of defining time ordered products to the problem of extending the distributions T_{n+1} defined by (169) from $M^{n+1} \setminus \Delta_{n+1}$ to all of M^{n+1} so that properties T1–T5 and T9 continue to hold for the extension. To find that extension, we now make

[7] Of course, if any T_{n+1} failed to satisfy any of these properties on $M^{n+1} \setminus \Delta_{n+1}$, we would have a proof that no definition of time ordered products could exist that satisfies T1–T9.

a Wick expansion of T_{n+1}, which follows from the Wick expansion at lower orders. That Wick expansion will contain c-number distribution coefficients, t, that are defined as distributions on a neighborhood of Δ_{n+1} in $M^{n+1} \setminus \Delta_{n+1}$. They possess a scaling expansion analogous to (162), with distributions u_k that are defined on $(\mathbb{R}^4)^n \setminus 0$. As we have just argued, time ordered products satisfying all of our conditions will exist if and only if the c-number distributions t defined away from Δ_{n+1} appearing in the Wick expansion for T_{n+1} analogous (159) can be extended to distributions defined on an open neighborhood of Δ_{n+1} in such a way that the distribution T_{n+1} defined by (159) continues to satisfy properties T1–T5. It is straightforward to check that this will be the case if and only if the extensions t satisfy the following five corresponding conditions:

t1 Locality/Covariance. The distributions $t = t_{j_1,\dots,j_{n+1}}$ are locally constructed from the metric in a covariant manner in the following sense. Let $\psi : M \to M'$ be a causality-preserving isometric embedding, so that $\psi^* g' = g$. Then eq. (160) holds for $m = n+1$.

t2 Scaling. The extended distributions t scale homogeneously up to logarithmic terms, in the sense that there is an $N \in \mathbb{N}$ such that (161) holds for $m = n+1$.

t3 Microlocal Spectrum Condition. The extension satisfies the wave front set condition that the restriction of $\mathrm{WF}(t)$ to the diagonal Δ_{n+1} is contained in $\{(x,k_1,\dots,x,k_{n+1}) \mid \sum k_i = 0\}$.

t4 Smoothness. t depends smoothly on the metric.

t5 Analyticity. For analytic spacetimes t depends analytically on the metric.

In summary, we have reduced the problem of defining time ordered products to the following question: Assume that time ordered products involving $\leq n$ factors have been constructed so as to satisfy our requirements T1–T9. Define T_{n+1} by (169) and define the distributions t on $M^{n+1} \setminus \Delta_{n+1}$ by the analogy of (159) for T_{n+1}, in a neighborhood of the diagonal. Can each t be extended to a distribution defined on a neighborhood of Δ_{n+1} so as to satisfy requirements t1–t5?

The answer to this question is "yes," and we shall now show how the desired extension of $t(x_1,\dots,x_{n+1})$ may be found. The idea is that, since the remainder in the scaling expansion (162) for t has an arbitrary low scaling degree for sufficiently large m by item (v), it can be extended to the diagonal Δ_{n+1} by continuity [18], i.e., there is no need to "renormalize" the remainder for sufficiently large but finite S. In fact, by Thm. 5.3 of [18], it is sufficient to choose any $S \geq d - 4n$ for this purpose. Furthermore, each term in the sum in the scaling expansion (162) can be written as $C^k(y) \cdot u_k(\xi_1,\dots,\xi_n)$ by (i). Each u_k is an almost homogeneous, Lorentz invariant n-point distribution on $(\mathbb{R}^4)^n \setminus 0$. As we will see presently in lemma 6 [63], this Minkowski distribution can be extended to a distribution on $(\mathbb{R}^4)^n$ with the same properties

[possibly with a higher N than that appearing (165)], by techniques in Minkowski space. It is this step that corresponds to the renormalization. As a consequence of the properties satisfied by the extension u, the corresponding extension t can be seen to satisfy t1)—t5), thus solving the renormalization problem for the time ordered products $T_{n+1}(\otimes_{i=1}^{n+1}\phi^{k_i}(x_i))$ with $n+1$ factors.

Lemma 6. Let $u \equiv u_{\mu_1\dots\mu_l}(\xi_1,\dots,\xi_n)$ be a Lorentz invariant tensor-valued distribution on $\mathbb{R}^{4n}\setminus 0$ which scales almost homogeneously with degree $\rho \in \mathbb{C}$ under coordinate rescalings, i.e.,

$$S_\rho^N u = 0 \quad \text{for some natural number } N. \tag{170}$$

where

$$S_\rho = \sum_{i=1}^n \xi_i^\mu \partial/\partial\xi_i^\mu + \rho\,. \tag{171}$$

Then u has a Lorentz invariant extension, also denoted u, to a distribution on $\mathbb{R}^{4n}$ which also scales almost homogeneously with degree ρ under rescalings of the coordinates. Moreover:

1. If $\rho \in \mathbb{Z}$, $\rho < 4n$, then u can be extended by continuity, the extension is unique, and $S_\rho^N u = 0$.

2. If $\rho \in \mathbb{C}\setminus\mathbb{Z}$ then the extension is unique, and $S_\rho^N u = 0$.

3. If $\rho \in \mathbb{Z}$, $\rho \geq 4n$, then the extension is not unique, and $S_\rho^{N+1} u = 0$. Two different extensions can differ at most by a distribution of the form $L\delta$, where L is a Lorentz-invariant partial differential operator in $\xi_1,\dots,\xi_n$ containing derivatives of degree $\rho - 4n$.

Proof: The proof of the lemma shows how the desired extension u can be constructed. We will first construct an extension that satisfies the almost homogeneous scaling property. This extension need not satisfy the Lorentz invariance properties. However, we will show that the extension can be modified, if necessary, so that the desired Lorentz-invariance property is satisfied, while retaining the desired almost homogeneous scaling behavior. The proof of the theorem given here differs from that given in [63], and thereby provides an alternative construction of the extension. A less general result of a similar nature for distributions with an *exactly* homogeneous scaling has previously been obtained in [72, Thms. 3.2.3 and 3.2.4]. Thus, our theorem generalizes this result to the case of almost homogeneous scaling. To simplify the notation, we set $x = (\xi_1,\dots,\xi_n) \in \mathbb{R}^{4n}$ throughout this proof.

The almost homogeneous scaling property of u, eq. (202), or the equivalent form of this condition (164) implies that $u(rx)$ can be written in the form

$$u(rx) = r^{-\rho}\sum_{k=0}^{N-1}\frac{(\log r)^k}{k!}v_k(x) \qquad r > 0\,, \tag{172}$$

where v_k are the distributions defined on $\mathbb{R}^{4n}\setminus 0$ by

$$v_k = S_\rho^k u\,. \tag{173}$$

Choose an arbitrary compact $4n-1$-dimensional surface $\Sigma \subset \mathbb{R}^{4n}$ homeomorphic to the sphere S^{4n-1} around the origin of $\mathbb{R}^{4n}$ that intersects each orbit of the scaling map $x \mapsto \mu x$ transversally and precisely once[8]. The first aim is to show that the distributions v_k can be restricted to Σ. To prove this, it is convenient to use the methods of microlocal analysis, in particular the following result [72]: If φ is a distribution on a manifold X with a submanifold Y, then φ can be restricted to Y if its wave front set (see Appendix C) satisfies $\mathrm{WF}(\varphi)|_Y \cap N^*Y = \emptyset$, where N^*Y is the "conormal bundle," defined as

$$N^*Y = \{(y,k) \in T_y^*X;\, y \in Y, k_i w^i = 0 \;\forall w \in T_yY\}\,. \tag{174}$$

We would like to apply this result to the situation $\Sigma = Y, \mathbb{R}^{4n} \setminus 0 = X$, and $v_k = \varphi$. To estimate the wave front set of the distributions v_k, we use another result from microlocal analysis [72]. Suppose A is a differential operator on X such that $A\varphi$ is smooth. Then $\mathrm{WF}(\varphi) \subset \mathrm{char}(A) \setminus 0$, where the characteristic set of A is defined by $\mathrm{char}(A) = \{(x,k) \in T_x^*X;\, a(x,k) = 0\}$, where a is the principal symbol of A. In our case, we have $S_\rho^{N-k} v_k = 0$, so

$$\mathrm{WF}(v_k) \subset \mathrm{char}(S_\rho^{N-k}) \setminus 0 = \left\{(x,k) \in T^*\mathbb{R}^{4n};\, \sum_i \xi_i \cdot k_i = 0, k \neq 0\right\} \tag{175}$$

because the principal symbol of S_ρ is given by $s(x,k) = \sum \xi_i \cdot k_i$, where we recall the notation $x = (\xi_1, \dots, \xi_n)$, and where we have set $k = (k_1, \dots, k_n) \in (\mathbb{R}^{4n})^*$. Assume now that $(x,k) \in N^*\Sigma$, and at the same time $(x,k) \in \mathrm{WF}(v_k)|_\Sigma$. Then, from the first condition, we have $w \cdot k = 0$ for all $w \in T_x\mathbb{R}^{4n}$ that are tangent to S, while from the second condition, we have $x \cdot k = 0$ and $k \neq 0$. Since Σ is transverse to the scaling orbits, it follows that $k = 0$, a contradiction. Hence $\mathrm{WF}(v_k)|_\Sigma \cap N^*\Sigma = \emptyset$, and v_k can be restricted to Σ. We denote points in Σ by $\hat{x}$, and we denote the restriction simply by $v_k(\hat{x})$, by the usual abuse of notation.

Let $\Sigma \subset \mathbb{R}^{4n}$ a submanifold of dimension $4n-1$ as above, and define, for $r > 0$

$$\Sigma_r = \{r\hat{x} \in \mathbb{R}^{4n};\, \hat{x} \in \Sigma\}\,. \tag{176}$$

We let $d^{4n}x$ be the usual $4n$-form on $\mathbb{R}^{4n}$ with the orientations induced from $\mathbb{R}^4$, i.e.,

$$d^{4n}x = d^4\xi_1 \wedge \cdots \wedge d^4\xi_n, \quad d^4\xi = d\xi^0 \wedge \cdots \wedge d\xi^3\,, \tag{177}$$

where we have put again $x = (\xi_1, \dots, \xi_n)$ to lighten the notation. We also define the 3-form w on $\mathbb{R}^4$ and the $4n-1$ form Ω on $\mathbb{R}^{4n}$ by

$$w(\xi) = \sum_{\mu=0}^{3} \xi^\mu d\xi_1 \wedge \dots \widehat{d\xi^\mu} \wedge \dots d\xi^3, \tag{178}$$

$$\Omega(x) = \sum_{i=1}^{n} d^4\xi_1 \wedge \dots w(\xi_i) \wedge \dots d^4\xi_n \tag{179}$$

[8]For example, we may choose Σ to be the sphere S^{4n-1} defined relative to some auxiliary Euclidean metric on $\mathbb{R}^{4n}$.

where a caret denotes omission. Because we are assuming that the surface Σ is transverse to the orbits of dilations in $\mathbb{R}^{4n}$, the map $(r,\hat{x}) \in \mathbb{R}_+ \times \Sigma \mapsto r\hat{x} \in \mathbb{R}^{4n} \setminus 0$ is an diffeomorphism. If $i_r : \Sigma_r \to \mathbb{R}^{4n}$ is the natural inclusion, then we may write

$$d^{4n}x = \frac{dr}{r} \wedge i_r^* \Omega\,. \tag{180}$$

Now let f be a testfunction of compact support on $\mathbb{R}^{4n} \setminus 0$, i.e., f is smooth, vanishes outside a compact set, and vanishes in an open neighborhood of 0. From the equation for $d^{4n}x$, and from eq. (172), we then get the following representation for $u(f)$:

$$\begin{aligned} u(f) &= \int_{\mathbb{R}^{4n}} u(x)\, f(x)\, d^{4n}x \\ &= \int_0^\infty \left(\int_{\Sigma_r} u(x)\, f(x) \Omega(x) \right) \frac{dr}{r} \\ &= \int_0^\infty r^{4n-1} \left(\int_{\Sigma_1} u(rx)\, f(rx) \Omega(x) \right) dr \\ &= \int_0^\infty \sum_{k=0}^{N-1} r^{4n-1-\rho} \frac{(\log r)^k}{k!} \left(\int_\Sigma v_k(x)\, f(rx)\, \Omega(x) \right) dr\,. \end{aligned} \tag{181}$$

The terms in the sum may be written as residue using the equality

$$r^a = \sum_k \frac{a^k (\log r)^k}{k!}\,. \tag{182}$$

For this, let $f_r(\hat{x})$ be the function on Σ defined by $f(r\hat{x})$. Then we may write

$$v_k(f_r) = \int_\Sigma v_k(x)\, f_r(x)\, \Omega(x)\,, \tag{183}$$

and we have

$$u(f) = \mathrm{Res}_{a=0} \sum_{k=0}^{N-1} \frac{1}{a^{k+1}} \int_0^\infty r^{a+4n-1-\rho} v_k(f_r)\, dr\,, \tag{184}$$

This formula is well defined because, since the support of f is bounded away from the origin in $\mathbb{R}^{4n}$, the distribution $r \mapsto v_k(f_r)$ is in fact a smooth test function on $\mathbb{R}_+$ whose support is compact and bounded away from $r = 0$. We would like to define the desired extension u' by generalizing formula (184) to arbitrary test functions f on $\mathbb{R}^{4n}$ whose support is not necessarily bounded away from the origin. If f is an arbitrary test function then $r \mapsto v_k(f_r)$ vanishes for sufficiently large $r > r_0$, but it no longer vanishes near $r = 0$. In that case, it is not obvious that the right side of (184) is still well-defined, and if so, whether it defines a meromorphic function of a. To show this, we let

$$h_k(r) := v_k(f_r) = \int_\Sigma v_k(x)\, f(rx)\, \Omega(x)\,, \tag{185}$$

and we observe that the distribution u' defined by

$$u'(f) := \mathrm{Res}_{a=0} \sum_{k=0}^{N-1} \frac{1}{a^{k+1}} \int_0^1 r^{a-\rho+4n-1} \Big(h_k(r) - \sum_{j=0}^{m} \frac{r^j}{j!} \frac{d^j h_k(0)}{dr^j} \Big) \, dr + \mathrm{Res}_{a=0} \sum_{k=0}^{N-1} \frac{1}{a^{k+1}} \int_1^\infty r^{a-\rho+4n-1} h_k(r) \, dr \tag{186}$$

is well-defined for all test functions f if m is chosen to be the largest integer $\le \mathrm{Re}\,\rho - 4n$. Indeed, the first integral on the right side is well defined and analytic for $\mathrm{Re}\,a > -1$, and the second term is well defined and analytic for all $a \in \mathbb{C}$. Thus, the terms on the right side of the above equation are linear functionals on the space of test functions that are meromorphic in a. Furthermore, if f has its support away from 0, then $h_k(r) = 0$ in an open neighborhood of $r = 0$, and we have $u'(f) = u(f)$. Finally, it can be shown using the methods described in chapter I, paragraph 3 of [53] that $u'(f)$ is not just a linear functional on the space of test-functions, but defines in fact a distribution on $\mathbb{R}^{4n}$. Consequently, (186) defines an extension u' of the distribution u.

We next need to analyze the scaling behavior of our extension u'. A straightforward calculation using eq. (186) shows that

$$(S^N_\rho u')(f) = -\mathrm{Res}_{a=0} \left\{ \frac{\partial^N}{\partial(\log \mu)^N} \sum_{k=1}^{N-1} \frac{\mu^a}{a^{k+1}} \left[\frac{r^{4n-\rho+a} h_k(0)}{4n-\rho+a} + \cdots + \frac{r^{4n-\rho+a+m} \frac{d^m}{dr^m} h_k(0)}{m!(4n-\rho+a+m)} \right]_{r=1}^{r=\mu} \right\}_{\mu=1} . \tag{187}$$

If we now assume that we are in case 3) of the lemma, i.e., $\rho \in \mathbb{N}_0 + 4n$, then $m = \rho - 4n$, and the expression evaluates to

$$(S^N_\rho u')(f) = \frac{\frac{d^{\rho-4n}}{dr^{\rho-4n}} h_{N-1}(0)}{(\rho-4n)!} . \tag{188}$$

The terms on the right side can be evaluated as follows using the definition of $h_{N-1}(r)$ and $v_{N-1}(x)$, see eqs. (185) and (173):

$$\frac{d^{\rho-4n}}{dr^{\rho-4n}} h_{N-1}(0) = \sum_{|\alpha|=\rho-4n} \left(\int_\Sigma x^\alpha S^{N-1}_\rho u(x) \, \Omega(x) \right) (\partial_\alpha f)(0) , \tag{189}$$

where $\alpha = (\alpha_1, \dots, \alpha_{4n}) \in \mathbb{N}_0^{4n}$ is a multi-index, and we are using the usual multi-index notation

$$\partial_\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_{4n}^{\alpha_{4n}}}, \quad |\alpha| = \sum_i \alpha_i, \quad x^\alpha = x_1^{\alpha_1} \cdots x_{4n}^{\alpha_{4n}} . \tag{190}$$

Alternatively, we may write

$$S^N_\rho u'(x) = \sum_{|\alpha|=\rho-4n} c^\alpha \partial_\alpha \delta(x) \tag{191}$$

in terms of the usual δ-function on $\mathbb{R}^{4n}$ concentrated at the origin. The numerical constants $c^\alpha \in \mathbb{C}$ are given by the formula

$$c^\alpha = \int_\Sigma F^\alpha(x)\,, \tag{192}$$

with F^α the (distributional) $(4n-1)-forms$ on Σ defined by

$$F^\alpha(x) := \frac{(-1)^{\rho-4n}}{(\rho-4n)!} x^\alpha S_\rho^{N-1} u(x)\cdot\Omega(x) \quad \in \mathcal{D}'\big(\Sigma; \wedge^{4n-1} T^*\Sigma\big)\,. \tag{193}$$

Since the delta-function is a homogeneous distribution of degree $-4n$, we have $S_\rho \partial_\alpha \delta = \partial_\alpha S_{4n} \delta = 0$, and therefore $S_\rho^{N+1} u' = 0$ by eq. (203). Thus our extension u' is again an almost homogeneous distribution.

One may repeat this argument also for the first and second case of the lemma. In those cases, one finds $S_\rho^N \hat{u} = 0$. Thus, summarizing, eq. (186) defines a distributional extension u' of u that is almost homogeneous. To simplify the notation, we will from now on denote this extension again by u.

We now investigate the Lorentz transformation properties of u. Our construction of the extension u given above involved a choice of a suitable Σ transverse to the orbits of the dilations. Since no Σ with the above properties exists that is at the same time invariant under the Lorentz group, the extension u just constructed will in general fail to be Lorentz invariant. Restoring the tensor indices on u, we find by a calculation using eq. (186) that for any test function $f \in C_0^\infty(\mathbb{R}^{4n})$ and any Lorentz transformation, Λ, we have

$$u_{\mu_1\dots\mu_l}(f) - \Lambda^{\nu_1}_{\mu_1}\dots\Lambda^{\nu_l}_{\mu_l} u_{\nu_1\dots\nu_l}(R(\Lambda)f) = \sum_{|\alpha|\le\rho-4n} b^\alpha_{\mu_1\dots\mu_l}(\Lambda)\partial_\alpha\delta(f), \tag{194}$$

where $(R(\Lambda)f)(x) = f(\Lambda^{-1}x)$ and the $b^\alpha_{\mu_1\dots\mu_l}(\Lambda)$ are complex constants, which would vanish if and only if the distribution u were Lorentz invariant. We now apply the differential operator S_ρ^{N+1} to both sides of the above equation. Since S_ρ is itself a Lorentz invariant operator, we have $R(\Lambda)S_\rho = S_\rho R(\Lambda)$. Therefore, since $S_\rho^{N+1} u = 0$, the operator S_ρ^{N+1} annihilates the left side of eq. (194), so we obtain

$$0 = S_\rho^{N+1} \sum_{|\alpha|\le \mathrm{Re}(\rho)-4n} b^\alpha_{\mu_1\dots\mu_l}(\Lambda)\partial_\alpha\delta = \sum_{|\alpha|\le \mathrm{Re}(\rho)-4n} (\rho-4n-|\alpha|)^{N+1} b^\alpha_{\mu_1\dots\mu_l}(\Lambda)\partial_\alpha\delta. \tag{195}$$

It follows immediately that $b^\alpha_{\mu_1\dots\mu_l}(\Lambda) = 0$, except possibly when $|\alpha| = \rho - 4n$, which evidently can only happen when ρ is an integer. Thus, focussing on that case, we have

$$u_{\mu_1\dots\mu_l}(f) - \Lambda^{\nu_1}_{\mu_1}\dots\Lambda^{\nu_l}_{\mu_l} u_{\nu_1\dots\nu_l}(R(\Lambda)f) = b^{\nu_1\dots\nu_{\rho-4n}}_{\mu_1\dots\mu_l}(\Lambda)\partial_{\nu_1}\dots\partial_{\nu_{\rho-4n}}\delta(f) \tag{196}$$

for all f and all Lorentz-transformations Λ. Using this equation, one finds the following transformation property for $b(\Lambda)$,

$$0 = b(\Lambda_1\Lambda_2) - b(\Lambda_1) - \mathrm{D}(\Lambda_1)b(\Lambda_2) \equiv (\delta b)(\Lambda_1, \Lambda_2), \tag{197}$$

where we have now dropped the tensor-indices and where D denotes the tensor representation of the Lorentz-group on the space $\mathrm{D} = (\otimes^l \mathbb{R}^4)^* \otimes (\otimes^{\rho-4n} \mathbb{R}^4)$. This relation is of cohomological nature. To see its relation to cohomology, one defines the following group-cohomology rings, see e.g. [59]:

Definition 3.1. Let G be a group, D a representation of G on a vector space V, and let c^n be the space of functionals $\xi_n : G^{\times n} \to V$. Let $\delta : c^n \to c^{n+1}$ be defined by

$$\begin{aligned}(\delta\xi_n)(g_1,\dots,g_{n+1}) &= \mathrm{D}(g_1)\xi_n(g_2,\dots,g_{n+1}) + \sum_{i=1}^{n}(-1)^i \xi_n(g_1,\dots,g_i g_{i+1},\dots,g_{n+1}) \\ &\quad + (-1)^{n+1}\xi_n(g_1,\dots,g_n)\,. \end{aligned} \tag{198}$$

Then $\delta^2 = 0$. The corresponding cohomology rings are defined as

$$H^n(G;\mathrm{D}) = \frac{\{\mathrm{Kernel}\,\delta : c^n \to c^{n+1}\}}{\{\mathrm{Image}\,\delta : c^{n-1} \to c^n\}}\,. \tag{199}$$

According to this definition, eq. (197) may be viewed [92] as saying that $b \in H^1(SO_0(3,1);\mathrm{D})$. It is a classical result of Wigner [111] that this ring is trivial for the Lorentz group and any finite-dimensional D. It follows that there is an a such that $b = \delta a$, or

$$b(\Lambda) = (\delta a)(\Lambda) \equiv a - \mathrm{D}(\Lambda)a \quad \forall \Lambda, \tag{200}$$

where a is an element in $H^0(SO_0(3,1);\mathrm{D}) = \mathrm{D} = (\otimes^l \mathbb{R}^4)^* \otimes (\otimes^{\rho-4n}\mathbb{R}^4)$. This enables us to define a modified extension $\hat{u}$ by

$$u'_{\mu_1\dots\mu_l} := u_{\mu_1\dots\mu_l} - a^{\nu_1\dots\nu_{\rho-4n}}_{\mu_1\dots\mu_l}\partial_{\nu_1}\dots\partial_{\nu_{\rho-4n}}\delta, \tag{201}$$

where we have now restored the tensor indices. It is easily checked that u' is Lorentz invariant and satisfies $S_\rho^{N+1}u' = 0$. In cases 1) and 2), u actually even satisfies $S_\rho^N u = 0$, so the modified extension (201) even satisfies $S_\rho^N u' = 0$. We have therefore accomplished the goal of constructing the desired extension of u in cases 1), 2) and 3).

The uniqueness statement immediately follows from the fact that the difference between any two extensions has to be a Lorentz-invariant derivative of the delta-function, $L\delta$, such that $S_\rho^{N+1}L\delta = 0$. Thus, L can be non-zero only when ρ is an integer, and L must have degree of precisely $\rho - 4n$. □

From the proof of the lemma, we get the following interesting proposition:

Proposition 1: Let $u(x)$ be a Lorentz invariant (possibly tensor-valued) distribution on $\mathbb{R}^{4n} \setminus 0$ which scales almost homogeneously with degree $\rho \in 4n + \mathbb{N}_0$ under coordinate rescalings, i.e.,

$$S_\rho^N u(x) = 0 \quad \text{for some natural number } N,\ x \neq 0, \tag{202}$$

Then u has a Lorentz invariant extension, also denoted u, to a distribution on $\mathbb{R}^{4n}$ which also scales almost homogeneously with degree ρ under rescalings of the coordinates. We have $S_\rho^{N+1} u = 0$, and

$$S_\rho^N u(x) = \sum_{|\alpha| = \rho - 4n} c^\alpha \partial_\alpha \delta(x) \tag{203}$$

in terms of the usual δ-function on $\mathbb{R}^{4n}$ concentrated at the origin. The numerical constants $c^\alpha \in \mathbb{C}$ are Lorentz-invariants, and are given by the formula

$$c^\alpha = \int_\Sigma F^\alpha(x) \,, \tag{204}$$

where $\Sigma \subset \mathbb{R}^{4n}$ is *any* closed $(4n-1)$ submanifold enclosing the origin $0 \in \mathbb{R}^{4n}$ which is transverse to the orbits of to the dilations of $\mathbb{R}^{4n}$. Here, the distributional $(4n-1)$-forms $F^\alpha \in \mathcal{D}'(\Sigma; \wedge^{4n-1} T^*\Sigma)$ on Σ are defined in eq. (193), and are closed,

$$dF^\alpha = 0 \,. \tag{205}$$

Proof: We only need to show that the $(4n-1)$-forms F^α are closed, and that the c^α are Lorentz invariants. We first compute

$$d\Omega(x) = 4n \, d^{4n} x \tag{206}$$

using the definition of the $(4n-1)$-form Ω, see eq. (178). By a straightforward computation using the definition of Ω, we also have

$$d[x^\alpha S_\rho^{N-1} u(x)] \wedge \Omega(x) = x^\alpha (S_0 + |\alpha|) [(S_0 - \rho)^{N-1} u(x)] \, d^{4n} x \,. \tag{207}$$

Using next the fact that $|\alpha| = \rho - 4n$, and that $S_\rho^N u = (S_0 - \rho)^N u = 0$, we find

$$d[x^\alpha S_\rho^{N-1} u(x)] \wedge \Omega(x) = -4n \, x^\alpha S_\rho^{N-1} u(x) \, d^{4n} x \tag{208}$$

so

$$\begin{aligned} dF^\alpha(x) &= \frac{(-1)^{\rho-4n}}{(\rho-4n)!} \, d[x^\alpha S_\rho^{N-1} u(x) \, \Omega(x)] && (209) \\ &= \frac{(-1)^{\rho-4n}}{(\rho-4n)!} \Big\{ d[x^\alpha S_\rho^{N-1} u(x)] \wedge \Omega(x) + x^\alpha S_\rho^{N-1} u(x) \, d\Omega(x) \Big\} = 0 \,. && (210) \end{aligned}$$

We would next like to show that c^α are Lorentz invariants, in the sense that $\Lambda^\alpha_\beta c^\beta = c^\alpha$ for any

Lorentz transformation. We have

$$\begin{aligned}
\Lambda^{\alpha}_{\beta} c^{\beta} &= \int_{\Sigma} \Lambda^{\alpha}_{\beta} F^{\beta}(x) \\
&= \int_{\Lambda^{*}\Sigma} \Lambda^{\alpha}_{\beta} F^{\beta}(\Lambda^{-1} x) \\
&= \int_{\Lambda^{*}\Sigma} F^{\alpha}(x) \\
&= \int_{\Sigma} F^{\alpha}(x) + \int_{U} dF^{\alpha}(x) \\
&= c^{\alpha}.
\end{aligned} \tag{211}$$

Here we have used in the first step the definition of c^{α}, in the second step we have used the standard transformation formula of an integral under a diffeomorphism, denoting by $\Lambda^{*}\Sigma$ image of Σ under the natural action of Λ on $\mathbb{R}^{4n}$. In the third step we have used that F^{α} itself is Lorentz invariant, and in the fourth step we have used Stoke's theorem for the open set $U \subset \mathbb{R}^{4n}$ such that $\partial U = -\Sigma \cup \Lambda^{*}\Sigma$, and in the fifth step we used $dF^{\alpha} = 0$. $\square$

In summary, we have now described how to construct the time ordered products $T_n(\otimes_{i=1}^{n} \phi^{k_i})$ of Wick monomials without derivatives. These construction can in principle be generalized to time ordered products of Wick monomials $\mathcal{O}_i$ containing derivatives by generalizing the Wick expansion to fields with derivatives. A non-trivial new renormalization condition now arises from T10, because S_0 contains derivatives. This condition is not automatically satisfied, but it is not difficult to see that we can change, if necessary, our construction of the time ordered products, so as to also satisfy T10 [66].

We finally have to consider condition T11. This condition is satisfied by our construction for T_1, but not in general for T_n when $n > 1$. The operational meaning of this requirement is that "derivatives can be freely pulled through the time-ordering symbol". This identity is a non-trivial requirement because both sides of the equation mean quite different things a priori: The first expression means the time ordered product of fields, one of which contains a total derivative, the second expression denotes the derivative, in the sense of distributions, of the algebra valued distribution given by the time ordered product of the fields without the total derivative. That these two quantities are actually the same is not obvious from the above construction, and is therefore an additional renormalization condition, called the "action Ward identity" in [34], and the "Leibniz rule" in [66]. It is shown in these two references how, starting from a prescription that satisfies T1—T10 but possibly does not satisfy this renormalization condition, one can go to a prescription which does.

The action Ward identity is at odds with conventions often found in standard textbooks on field theory in Minkowski spacetime [110], where the derivative is not taken to commute with T_n. To illustrate this difference in point of view, consider the time ordered product $T_2(\phi(x) \otimes \phi(y))$. According to condition T11, we have $(\Box_x - m^2) T_2(\phi(x) \otimes \phi(y)) = T_2((\Box_x - m^2)\phi(x) \otimes \phi(y))$. In our approach, the time ordered products need not vanish when acting on a factor of

the wave equation, so this quantity does not need to vanish. In fact, one can see that the time-ordered product under consideration is uniquely determined by the properties T1—T10, and we have $T_2((\Box_x - m^2)\phi(x) \otimes \phi(y)) = i\hbar\delta(x,y)\mathbb{1}$. In standard approaches, on the other hand, it is assumed that the time ordered product vanishes when acting on $(\Box_x - m^2)\phi(x)$, because the time-ordering symbol is viewed as on operation acting on on-shell quantized fields, rather than just classical polynomial expressions in **P**. On the other hand, in most standard approaches, it is not assumed that derivatives commute with T_2. In this way, one reaches the same conclusion for the example just considered, and both viewpoints are consistent for that example. However, the standard viewpoint gets very awkward in general when considering more complicated time ordered products of fields with derivatives, for a discussion see e.g. [35]. This is because it is in general inconsistent to assume that a time ordered product containing a factor $\mathcal{O}\Box\phi$ vanishes, because of possible anomalies. On the other hand, the Leibniz rule can always be satisfied, and possible anomalies can thereby be analyzed consistently.

3.4 Examples

Here we illustrate the above general construction of the time-ordered product by some simple examples. The simplest non-trivial example of a time ordered product with one factor is $T_1(\phi^2(x)) =: \phi^2(x) :_{\mathrm{H}}$. Using the definition of the locally normal ordered product, this may be viewed as a "point-splitting" definition, see e.g. [24]. Consider next the time ordered product $T_2(\phi^2(x_1) \otimes \phi^2(x_2))$. By T8, it is defined for non-coincident points $x_1 \neq x_2$ by the prescription

$$T_2(\phi^2(x_1) \otimes \phi^2(x_2)) = \begin{cases} :\phi^2(x_2):_{\mathrm{H}} \star_\hbar :\phi^2(x_1):_{\mathrm{H}} & \text{when } x_1 \notin J^+(x_2); \\ :\phi^2(x_1):_{\mathrm{H}} \star_\hbar :\phi^2(x_2):_{\mathrm{H}} & \text{when } x_1 \notin J^-(x_2). \end{cases} \tag{212}$$

In order to extend the definition to coincident points $x_1 = x_2$, i.e., to make the time-ordered product a well defined distribution on the entire product manifold M^2, we now use the expansion procedures described in general in the previous section. Using the definition of the product $\star_\hbar$, and of the locally normal ordered products, we have

$$:\phi^2(x_1):_{\mathrm{H}} \star_\hbar :\phi^2(x_2):_{\mathrm{H}} = \\ :\phi^2(x_1)\phi^2(x_2):_{\mathrm{H}} - 2\hbar H(x_1,x_2) :\phi(x_1)\phi(x_2):_{\mathrm{H}} + \hbar^2 H(x_1,x_2)^2\, \mathbb{1}\,, \tag{213}$$

for points x_1, x_2 that are sufficiently close to each other so that the local Hadamard parametrix $H(x_1,x_2)$ is well-defined. Using furthermore the definition of the local Feynman parametrix H_F (see eq. (520)) and

$$\vartheta(T(x) - T(y))\, H(x,y) + \vartheta(T(y) - T(x))\, H(y,x) = iH_F(x,y) \tag{214}$$

we can write the time ordered product under consideration as

$$T_2(\phi^2(x_1) \otimes \phi^2(x_2)) = \\ :\phi^2(x_1)\phi^2(x_2):_{\mathrm{H}} + 2(\hbar/i)H_F(x_1,x_2) :\phi(x_1)\phi(x_2):_{\mathrm{H}} + (\hbar/i)^2 H_F(x_1,x_2)^2\, \mathbb{1}\,, \tag{215}$$

for non-coinciding points x, y. This is the desired local Wick-expansion. Comparing with eq. (159), we read off

$$t_{0,0}(x_1,x_2) = 1, \quad t_{1,1}(x_1,x_2) = (\hbar/i)H_F(x_1,x_2), \quad t_{2,2}(x_1,x_2) = (\hbar/i)^2 H_F(x_1,x_2)^2 \tag{216}$$

for the coefficients in the Wick expansion. The coefficients $t_{0,0}, t_{1,1}$ may be extended to coincident points $x = y$ by continuity, because their scaling degree is 0 resp. 2, which is less than 4, but the distribution $t_{2,2}$ has scaling degree 4 and therefore cannot be extended to the diagonal by continuity, but must instead be extended non-trivially. Actually, since $t_{2,2}$ is the square of the distribution H_F with singularities on the lightcone, it is instructive to check explicitly that it is even defined for non-conincident points that are on the lightcone. This can be done using the wave front set: For $x_1 \notin J^+(x_2)$, the pair $(x_1,k_1;x_2,k_2) \in T^*(M^2)$ is in the wave front set of H_F (see appendix C) if and only if x_1 and x_2 can be joined by a null-geodesic $\gamma : (0,1) \to M$, with $\dot\gamma(0) = k_1$ and $\dot\gamma(1) = -k_2$, with $k_1 \in V_+^*$. Similarly, for $x_1 \notin J^-(x_2)$, the pair $(x_1,k_1;x_2,k_2) \in T^*(M^2)$ is in the wave front set if and only if x_1 and x_2 can be joined by a null-geodesic $\gamma : (0,1) \to M$, with $\dot\gamma(0) = k_1$ and $\dot\gamma(1) = -k_2$, with $k_1 \in V_-^*$. It follows that, when $x_1 \neq x_2$, elements $(x_1,k_1,x_2,k_2) \in \mathrm{WF}(H_F)$ can never add up to the zero element. Thus, by the general theorems about the wave front set summarized in appendix C, arbitrary powers $H_F(x_1,x_2)^n$ exist in the distributional sense, i.e., as distributions on $M^2 \setminus \Delta_2$. On the other hand, when $x_1 = x_2$, arbitrary elements of the form $(x_1,k,x_2,-k)$ are in $\mathrm{WF}(H_F)$. Thus, for coincident points, the elements in the wave front set can add up to zero, and the product $H_F(x_1,x_2)^n$ is therefore not defined as a distribution on all of M^2, i.e., including coincident points.

In order to extend $t_{2,2}$ to a well-defined distribution to all of M^2, we now need to perform the scaling expansion of $t_{2,2}$, which in turn can be obtained from the scaling expansion of H_F. The latter can be found using expansions for the recursively defined coefficients in the local Hadamard parametrix, see e.g. [24]. Up to numerical prefactors, it is given by (we assume for simplicity that $m^2 = 0$)

$$H_F(\exp_y \xi, y) \sim \frac{1}{\xi^2 + i0} + R_{\mu\nu}(y)\Big(-\frac{1}{6}\frac{\xi^\mu\xi^\nu}{\xi^2+i0} + \frac{1}{12}\eta^{\mu\nu}\log(\xi^2+i0)\Big) + \dots, \tag{217}$$

where the dots stand for a remainder with scaling degree < 2, where $\xi \in T_yM$ has been identified with a vector in $\mathbb{R}^4$ via a tetrad, and where $\xi^2 = \eta_{\mu\nu}\xi^\mu\xi^\nu$. From this we obtain the first terms in the scaling expansion of $t_{2,2}$ up to numerical prefactors as

$$t_{2,2}(\exp_y \xi, y) \sim u(\xi) + R_{\mu\nu}(y)\, u^{\mu\nu}(\xi) + \dots \tag{218}$$

where the dots stand for terms of scaling degree less than 2. The distributions u and $u^{\mu\nu}$ are defined on $\mathbb{R}^4 \setminus 0$ and is given there by

$$u(\xi) = \frac{1}{(\xi^2+i0)^2}, \quad u^{\mu\nu}(\xi) = -\frac{1}{3}\frac{\xi^\mu\xi^\nu}{(\xi^2+i0)^2} + \frac{1}{6}\frac{\eta^{\mu\nu}\log(\xi^2+i0)}{\xi^2+i0}. \tag{219}$$

u has scaling degree 4, while $u^{\mu\nu}$ has scaling degree 2. Thus, by lemma 6, we need to extend non-trivially only u, while $u^{\mu\nu}$ and the remainder (i.e., the dots in the scaling expansion of $t_{2,2}$) can be extended by continuity. An extension to all of $\mathbb{R}^4$ (i.e., including $\xi = 0$) of u can easily be guessed, but we here prefer to give a systematic method, which is needed anyway in more complicated examples. A constructive method to obtain an extension of u is provided by lemma 6. However, that has the disadvantage of being somewhat complicated because it involves a non-Lorentz invariant surface S at intermediate steps, which is awkward in concrete calculations[9]. Instead we here present a different method, that is more practical and works in a wide class of examples. That method is based upon the fact that, for complex scaling degree, there is a unique extension of a homogeneous distribution by lemma 6. The method has also appeared in the context of BPHZ-renormalization in momentum space under the name "analytic renormalization" [98, 99, 100].

Consider instead of u the distribution given by

$$u_a(\xi) = \frac{1}{(\xi^2 + i0)^{2-a}}, \quad a \in \mathbb{C} \setminus \mathbb{Z}. \tag{220}$$

By contrast to u, this is well defined on all of $\mathbb{R}^4$, see e.g. [53], and also [86] for a treatment of such so-called "Riesz-distributions". An extension u' of u can now be obtained by taking the residue of the meromorphic function $a \mapsto u_a(f)/a$,

$$u'(f) = \mathrm{Res}_{a=0} \frac{u_a(f)}{a}. \tag{221}$$

Indeed, if the support of f excludes 0, then $u'(f)$ obviously must coincide with $u(f)$, because we may then use formula (219) to get $u(f)$. The almost homogeneous scaling property of $u'(f)$ under rescalings of $f(\xi) \to f(\mu\xi)$ also immediately follows from the definition. To get a more explicit formula for the extension, we compute the fourier transform of u_a, given up to numerical factors by [86]

$$\hat{u}_a(p) = 4^a \frac{\Gamma(a)}{\Gamma(2-a)} (p^2 - i0)^{-a}. \tag{222}$$

We expand this expression around $a = 0$ using the well-known residue of the Γ-function at 0 and substitute the resulting expression into eq. (221). We obtain, up to numerical prefactors

$$\hat{u}'(p) = \ln[l^2(p^2 - i0)] \tag{223}$$

where l is some constant. Taking an inverse fourier transform then gives the desired extension

$$u'(\xi) = -\frac{1}{2}\partial^2 \left(\frac{\log[l^{-2}(\xi^2 + i0)]}{\xi^2 + i0} \right). \tag{224}$$

[9] Note, however, that this is not an obstacle in the corresponding "Euclidean situation", where one may take S simply to be a Euclidean sphere.

where $\partial^2 = \eta^{\mu\nu}\partial^2/\partial\xi^\mu\partial\xi^\nu$. Note that the extension has acquired a logarithm, which is a general phenomenon according to lemma 6. Different choices of l change the extension by a term proportional to $\delta^4(\xi)$, and thus correspond to the different extensions of $u(\xi)$. Thus, inserting this extension into the scaling expansion of $t_{2,2}$, we obtain the desired extension of $T_2(\phi^2(x_1)\otimes\phi^2(x_2))$.

Our last example is the time ordered product $T_3(\phi^3(x_1)\otimes\phi^3(x_2)\otimes\phi^4(x_3))$ with 3 factors. The terms in the Wick expansion of this quantity that need to be extended non-trivially from $M^3\setminus\Delta_3$ to M^3 are

$$t_{3,3,2}(x_1,x_2,x_3) = t_{1,1}(x_1,x_2)t_{1,1}(x_2,x_3)t_{2,2}(x_1,x_3), \tag{225}$$

$$t_{3,3,4}(x_1,x_2,x_3) = t_{1,1}(x_1,x_2)t_{2,2}(x_2,x_3)t_{2,2}(x_1,x_3). \tag{226}$$

All other terms are either already well-defined as distributions on all of M^3 (assuming the corresponding time ordered products with 2 factors have been defined), or can be extended by continuity. We focus on the last term $t_{3,3,4}$. Again, for the sake of illustration of the general construction, we first verify explicitly that this distribution is indeed well-defined on $M^3\setminus\Delta_3$. Consider a point $(x_1,x_2,x_3)\notin\Delta_3$. Then it must be possible to separate one point, from the remaining two points by a Cauchy surface. For definiteness, let us assume that this point is x_3, and that $x_1,x_2\notin J^+(x_3)$. Then $(x_1,k_1;x_3,k_3)$ is in the wave front set of $t_{2,2}(x_1,x_3)$ if and only if $k_1\sim -k_3$, and if $k_1\in V_+^*$. Likewise, $(x_2,p_2;x_3,p_3)$ is in the wave front set of $t_{2,2}(x_2,x_3)$ if and only if $p_2\sim -p_3$, and if $p_2\in V_+^*$. Finally $(x_1,q_1;x_2,q_2)$ is in the wave front set of $t_{1,1}(x_1,x_2)$ iff $q_1\sim -q_2$ and $q_1\in V_\pm^*$ when $x_1\notin J^\pm(x_2)$, or iff $q_1=-q_2$ when $x_1=x_2$. We now add up these wave front set elements, viewed in the obvious way as elements in $T^*_{x_1}M\times T^*_{x_2}M\times T^*_{x_3}M$. We obtain the set

$$S=\{(x_1,k_1+q_1;x_2,p_2+q_2;x_3,k_3+p_3)\}. \tag{227}$$

Assume first that $x_1=x_2$. Clearly, if e.g. $k_1+q_1=0$, then $q_1\in V_-^*$, so $p_2+q_2=p_2-q_1\neq 0$, because $p_2\in V_+^*$. Thus, S cannot contain the zero element, and the product defining $t_{3,3,4}$ is well-defined near (x_1,x_2,x_3) by thm. 5. Similarly, if $x_1\notin J^-(x_2)$, then $q_2\in V_+^*$, and again $p_2+q_2\neq 0$, and again, S cannot contain the zero element. The same type of argument can be made for all other configurations of the points, except the configuration $x_1=x_2=x_3$. Thus, by the general existence theorem 5 for products of distributions, $t_{3,3,4}$ is indeed well-defined as a distribution on $M^3\setminus\Delta_3$.

We next would like to construct an extension of $t_{3,3,4}$ along the lines of our general construction. Thus, we must determine the scaling expansion of $t_{3,3,4}$. It can be obtained from the expansions of the (extended) distributions $t_{2,2}$ and of $t_{1,1}$ that were constructed above. We focus on the terms that require a non-trivial extension (up to numerical prefactors):

$$t_{3,3,4}(\exp_y\xi_1,\exp_y\xi_2,y)\sim u(\xi_1,\xi_2)+R_{\mu\nu}(y)u^{\mu\nu}(\xi_1,\xi_2)+R_{\mu\nu\sigma\rho}(y)u^{\mu\nu\sigma\rho}(\xi_1,\xi_2)+\dots, \tag{228}$$

where u is the distribution defined on $(\mathbb{R}^4)^2\setminus 0$ given by

$$u(\xi_1,\xi_2)=\frac{1}{4}\partial_1^2\Big(\frac{\log[l^{-2}(\xi_1^2+i0)]}{\xi_1^2+i0}\Big)\partial_2^2\Big(\frac{\log[l^{-2}(\xi_2^2+i0)]}{\xi_2^2+i0}\Big)\frac{1}{(\xi_1-\xi_2)^2+i0}$$

where $u^{\mu\nu}$ is the distribution defined on $(\mathbb{R}^4)^2 \setminus 0$ given by

$$u^{\mu\nu}(\xi_1,\xi_2) = \\ -\frac{1}{2}\partial_1^2\Big(\frac{\log[l^{-2}(\xi_1^2+i0)]}{\xi_1^2+i0}\Big)\Big(-\frac{1}{3}\frac{\xi_2^\mu\xi_2^\nu}{(\xi_2^2+i0)^2}+\frac{1}{6}\frac{\eta^{\mu\nu}\log[l^{-2}(\xi_2^2+i0)]}{\xi_2^2+i0}\Big)\frac{1}{(\xi_1-\xi_2)^2+i0}+(\xi_1\leftrightarrow\xi_2) \tag{229}$$

and where $u^{\mu\nu\sigma\rho}$ is the distribution on $(\mathbb{R}^4)^2 \setminus 0$ defined by

$$u^{\mu\nu\sigma\rho}(\xi_1,\xi_2) = \frac{1}{4}\partial_1^2\Big(\frac{\log[l^{-2}(\xi_1^2+i0)]}{\xi_1^2+i0}\Big)\partial_2^2\Big(\frac{\log[l^{-2}(\xi_2^2+i0)]}{\xi_2^2+i0}\Big) \\ \cdot\Big(-\frac{1}{6}\frac{\xi_1^\mu\xi_1^\sigma\xi_2^\nu\xi_2^\rho}{[(\xi_1-\xi_2)^2+i0]^2}-\frac{1}{12}\frac{\eta^{\mu\sigma}(\xi_1^\nu\xi_2^\rho+2\xi_1^\nu\xi_1^\rho)}{(\xi_1-\xi_2)^2+i0}+\frac{1}{24}\eta^{\mu\sigma}\eta^{\nu\rho}\log\{l^{-2}[(\xi_1-\xi_2)^2+i0]\}\Big) \\ +(\xi_1\leftrightarrow\xi_2) \tag{230}$$

The dots in eq. (228) again represent a remainder. This now has scaling degree 6 and can thus be extended by continuity, while the 3 terms in the scaling expansion that are explicitly given have scaling degree 10 for the first term respectively 8 for the second and third term. They must thus be extended non-trivially. The extension of the corresponding distributions $u, u^{\mu\nu}, u^{\mu\nu\sigma\rho}$ now can no longer be found by trial and error, but one must use a constructive method, such as that given in the proof of lemma 6. We will again not use this method here, but instead use a variant of the method given above. For this, we consider the distribution

$$u_{a,b,c}(\xi_1,\xi_2) = \frac{1}{(\xi_1^2+i0)^{2-a}(\xi_2^2+i0)^{2-b}[(\xi_1-\xi_2)^2+i0]^{2-c}}\,. \tag{231}$$

It can be checked using wave-front arguments similar to that given above that this distributional product is well-defined on $(\mathbb{R}^4)^2 \setminus 0$ for $a,b,c \in \mathbb{C}\setminus\mathbb{Z}$. Furthermore, by Lemma 6, if $a+b+c \notin \mathbb{Z}$ this distribution has a unique extension to all of $(\mathbb{R}^4)^2$. We define the desired extension of u by the expression

$$u'(f) = \mathrm{Res}_{c=1}\mathrm{Res}_{b=0}\mathrm{Res}_{a=0}\,\frac{u_{a,b,c}(f)}{ab(c-1)}\,. \tag{232}$$

This is an extension, because one can check that $u'(f)$ conicides with $u(f)$ for any f whose support excludes $\xi_1 = \xi_2 = 0$, and it is also clearly Lorentz invariant and has the desired almost homogeneous scaling behavior. To get a more explicit expression for u', we perform a fourier transformation of $u_{a,b,c}$ using eq. (222) and eq. (23) of [23]. This gives, up to numerical factors

$$\hat{u}_{a,b,c}(p_1,p_2) = \frac{4^{a+b+c}}{\Gamma(4-a-b-c)\Gamma(2-a)\Gamma(2-b)\Gamma(2-c)}I_{a,b,c}(p_1,p_2) \tag{233}$$

where

$$
\begin{aligned}
&I_{a,b,c}(p_1,p_2) = \\
&[(p_1+p_2)^2 - i0]^{2-a-b-c}\Gamma(c)\Gamma(a+b+c-2)\Gamma(2-a-c)\Gamma(2-c-b)\times \\
&F_4\left(c, a+b+c-2, a+c-1, b+c-1 \middle| \frac{p_1^2}{(p_1+p_2)^2-i0}, \frac{p_2^2}{(p_1+p_2)^2-i0}\right) + \\
&[(p_1+p_2)^2 - i0]^{-a}(p_2^2-i0)^{2-b-c}\Gamma(a)\Gamma(2-b)\Gamma(2-a-c)\Gamma(b+c-2)\times \\
&F_4\left(a, 2-b, a+c-1, 3-b-c \middle| \frac{p_1^2}{(p_1+p_2)^2-i0}, \frac{p_2^2}{(p_1+p_2)^2-i0}\right) + \\
&[(p_1+p_2)^2 - i0]^{-b}(p_1^2-i0)^{2-a-c}\Gamma(b)\Gamma(2-a)\Gamma(a+c-2)\Gamma(2-c-b)\times \\
&F_4\left(b, 2-a, 3-a-c, b+c-1 \middle| \frac{p_1^2}{(p_1+p_2)^2-i0}, \frac{p_2^2}{(p_1+p_2)^2-i0}\right) + \\
&[(p_1+p_2)^2 - i0]^{c-2}(p_1^2-i0)^{2-a-c}(p_2^2-i0)^{2-b-c}\times \\
&\Gamma(4-a-b-c)\Gamma(2-c)\Gamma(a+b-2)\Gamma(b+c-2)\times \\
&F_4\left(4-a-b-c, 2-c, 3-a-c, 3-b-c \middle| \frac{p_1^2}{(p_1+p_2)^2-i0}, \frac{p_2^2}{(p_1+p_2)^2-i0}\right).
\end{aligned}
$$

Here, F_4 is the Appell function, defined by

$$
F_4(\alpha,\beta,\gamma,\delta|z_1,z_2) = \sum_{j_1,j_2=0}^{\infty} \frac{(\alpha)_{j_1+j_2}(\beta)_{j_1+j_2}}{(\gamma)_{j_1}(\delta)_{j_2}} z_1^{j_1} z_2^{j_2}, \tag{234}
$$

with $(\alpha)_j$ the Pochhammer symbol. The fourier transform of the extension is then given by

$$
\hat{u}'(p_1,p_2) = \mathrm{Res}_{c=1}\mathrm{Res}_{b=0}\mathrm{Res}_{a=0}\frac{\hat{u}_{a,b,c}(p_1,p_2)}{ab(c-1)}, \tag{235}
$$

which may be evaluated readily using the Laurent expansion of the Gamma-function. It is worth noting that the extension u' given by expression (232) now implicitly contains third powers of the logarithm, thus again confirming the general theorem that there are logarithmic corrections to the naively expected homogeneous scaling behavior.

3.5 Ghost fields and vector fields

The above algebraic construction of Wick-powers and their time-ordered products may be generalized to a multiplet of scalar or tensor fields satisfying a system of wave equations on M with local covariant coefficients or to Grassmann valued fields. In the BRST approach to gauge theory, the relevant fields are (gauge fixed) vector fields, and ghost fields.

Classical ghost fields are valued in the Grassmann algebra E. For gauge theory, the relevant ghost fields are described, at the free level, by the Lagrangian

$$
\mathbf{L}_0 = -id\bar{C} \wedge *dC. \tag{236}
$$

The fields $C, \bar{C}$ are independent and take values in the Grassmann algebra E. In particular, the "bar" over $\bar{C}$ is a purely conventional notation and is *not* intended to mean any kind of conjugation. The non-commutative *-algebra $\mathbf{W}_0$ corresponding to this classical Lagrangian is described as follows. As above, we consider a bi-distribution $\omega^{\rm s}(x,y)$ on $M \times M$ of Hadamard form (we put a superscript "s" for "scalar"), and we consider distributions u on M^n which are *anti*-symmetric in the variables, and which satisfy the wave-front condition (126). With each such distribution, we associate a generator $F(u)$, which we (purely formally) write as

$$F(u) = \int u(x_1, \dots, x_n; y_1, \dots, y_m) : C(x_1) \cdots C(x_n) \bar{C}(y_1) \cdots \bar{C}(y_m) :_\omega \, dx_1 \dots dx_n \, dy_1 \dots dy_n \,. \quad (237)$$

We now define a $\star_\hbar$-product between such generators. This is again defined by eq. (121), where the derivative operator (124) is now given by

$$<\mathcal{D}> = -i \int \frac{\delta_L}{\delta C(x)} \omega^{\rm s}(x,y) \frac{\delta_R}{\delta \bar{C}(y)} - \frac{\delta_L}{\delta \bar{C}(x)} \omega^{\rm s}(x,y) \frac{\delta_R}{\delta C(y)} \, dxdy \,. \quad (238)$$

Here, as above, it is understood that a functional derivative acting on $F(u)$ is executed by formally treating the fields in the normal ordered expression as classical fields, i.e., by formally identifying $: C(x_1) \cdots C(x_n) \bar{C}(y_1) \cdots \bar{C}(y_m) :_\omega$ with the classical field expression. The operation * of conjugation is defined as $C(x)^* = C(x)$ and $\bar{C}(x)^* = \bar{C}(x)$. This is consistent with the product. It leads to the anti-commutation relations for the ghost fields,

$$\begin{aligned} \bar{C}(x) \star_\hbar C(y) + C(y) \star_\hbar \bar{C}(x) &= \hbar \Delta^{\rm s}(x,y) \mathbb{1} \,, && (239) \\ C(x) \star_\hbar C(y) + C(y) \star_\hbar C(x) &= \bar{C}(x) \star_\hbar \bar{C}(y) + \bar{C}(y) \star_\hbar \bar{C}(x) = 0 \,, && (240) \end{aligned}$$

where we have put a superscript on "s" the scalar causal propagator $\Delta^{\rm s}$ to distinguish it from the vector propagator below. The field equations may be implemented, as in the scalar case, by dividing $\mathbf{W}_0$ by the ideal $\mathcal{J}_0$ generated by $\Box C(x)$ and $\Box \bar{C}(x)$. Time-ordered products of Grassmann fields are also defined in the same way as above, the only minor difference being that they are not symmetric in the tensor factors, but have graded symmetry according to the Grassmann parity of the arguments. For example, T6 reads instead

$$T_n(\dots \otimes \mathcal{O}_1(x_j) \otimes \mathcal{O}_2(x_{j+1}) \otimes \dots) = (-1)^{\varepsilon_j \varepsilon_{j+1}} T_n(\dots \otimes \mathcal{O}_2(x_{j+1}) \otimes \mathcal{O}_1(x_j) \otimes \dots) \,. \quad (241)$$

There are similar signs also in T9.

We next consider 1-form (or vector) fields, A. In the Lorentz gauge, their classical dynamics is described by the Lagrangian

$$\mathbf{L}_0 = \frac{1}{2} (dA \wedge *dA + \delta A \wedge *\delta A) \,. \quad (242)$$

where $\delta = *d*$ is the co-differential (divergence). Their equation of motion is the canonical wave equation for vectors, $(d\delta + \delta d)A = 0$, or

$$(g_{\mu\nu}\Box + R_{\mu\nu})A^{\nu} = 0 \tag{243}$$

in component notation. It is seen from the component form of the equation that it is hyperbolic in nature, and hence has unique fundamental retarded and advanced solutions, Δ^{v}_{A} and Δ^{v}_{R}, where we have put a superscript "v" in order to distinguish them from their scalar counterparts.

To define the corresponding quantum algebra of observables, we proceed by analogy with the scalar case. For this, we pick an arbitrary distribution ω^{v} taking values in $T^*M \times T^*M$ of Hadamard form. Thus, $\omega^{\mathrm{v}}(x,y)$ satisfies the vector equations of motion (243) in x and y, its anti-symmetric part is given by $i\Delta^{\mathrm{v}}(x,y)$, where Δ^{v} is the difference between the fundamental advanced and retarded vector causal propagators, and its wave-front set is given by eq. (119). The algebra $\mathbf{W}_0$ is generated by expressions of the form

$$F(u) = \int u(x_1, \dots, x_n) :A(x_1) \dots A(x_n):_{\omega} dx_1 \dots dx_n\,, \tag{244}$$

where $u(x_1, \dots, x_n)$ is a distribution with wave front set (126), now taking values in the bundle $TM \times \cdots \times TM$, and the *-operation is declared by $A(x)^* = A(x)$. The $\star_\hbar$-product is again defined by eq. (121), where the derivative operator (124) is now given by

$$_{<}\mathcal{D}_{>} = \int \frac{\delta_L}{\delta A(x)} \omega^{\mathrm{v}}(x,y) \frac{\delta_R}{\delta A(y)} dxdy\,. \tag{245}$$

From this, we can calculate the commutation relations for the field $A(x) =: A(x):_{\omega}$,

$$A(x) \star_\hbar A(y) - A(y) \star_\hbar A(x) = i\hbar\Delta^{\mathrm{v}}(x,y)\, 1\!\!1\,. \tag{246}$$

The construction of Wick powers and their time-ordered products is completely analogous to the scalar case, the only difference is that the Hadamard scalar parametrix H must be replaced by a vector Hadamard parametrix, whose construction is described in Appendix C.2.

3.6 Renormalization ambiguities of the time-ordered products

In the previous section, we have described the construction of local and covariant renormalized time ordered products in globally hyperbolic Lorentzian curved spacetimes. We now address the issue to what extend the time ordered products are unique. Thus, suppose we are given two prescriptions, called $T = \{T_n\}$ and $\hat{T} = \{\hat{T}_n\}$, satisfying the conditions T1—T11. We would like to know how they can differ. To characterize the difference, we introduce a hierarchy $D = \{D_n\}$ of linear functionals with the following properties. Each D_n is a linear map

$$D_n : \mathbf{P}^{k_1}(M) \otimes \cdots \otimes \mathbf{P}^{k_n}(M) \to \mathbf{P}^{k_1/\dots/k_n}(M^n)[[\hbar]]\,, \tag{247}$$

where we denote by $\mathbf{P}^{k_1/\cdots/k_n}(M^n)$ the space of all distributional local, covariant functionals of ϕ and its covariant derivatives $\nabla^k\phi$, of m^2, of the metric, and of the Riemann tensor and its covariant derivatives $\nabla^k R$, which are supported on the total diagonal, and which take values in the bundle

$$\bigwedge^{k_1} T^*M \times \cdots \times \bigwedge^{k_n} T^*M \subset \bigwedge^{k_1+\cdots+k_n} T^*M^n \tag{248}$$

of antisymmetric tensors over M^n. Thus, if $\mathcal{O}_i \in \mathbf{P}^{k_i}(M)$, then $D_n(\otimes_i \mathcal{O}_i) \in \mathbf{P}^{k_1/\cdots/k_n}(M^n)$, and D_n is a (distributional) polynomial, local, covariant functional of ϕ, the mass, m^2, and the Riemann tensor and its derivatives taking values in the $k_1 + \cdots + k_n$ forms over M^n, which is supported on the total diagonal, i.e.,

$$\text{supp}\, D_n(\mathcal{O}_1(x_1) \otimes \cdots \otimes \mathcal{O}_n(x_n)) = \{x_1 = x_2 = \cdots = x_n\} = \Delta_n\,. \tag{249}$$

It is a k_1-form in the first variable x_1, a k_2-form in the second variable x_2, etc.

The difference between two prescriptions T and $\hat{T}$ for time ordered products satisfying T1—T11 may now be expressed in terms of a hierarchy $D = \{D_n\}$ as follows. Let $F = \int f \wedge \mathcal{O}$ be an integrated local functional $\mathcal{O} \in \mathbf{P}(M)$, and formally combine the time-ordered functionals into a generating functional written

$$T(\mathrm{e}_\otimes^F) := \sum_{n=0}^{\infty} \frac{1}{n!} T_n(F^{\otimes n}), \tag{250}$$

where $\exp_\otimes$ is the standard map from the vector space of local actions to the tensor algebra (i.e., the symmetric Fock space) over the space of local action functionals. We similarly write $D(\mathrm{e}_\otimes^F)$ for the corresponding generating functional obtained from D. The difference between the time-ordered products T and $\hat{T}$ may now be expressed in the following way [62]:

$$\hat{T}\left(\mathrm{e}_\otimes^{iF/\hbar}\right) = T\left(\mathrm{e}_\otimes^{i[F+D(\exp_\otimes F)]/\hbar}\right). \tag{251}$$

where $D = \{D_n\}$ is a hierarchy of functionals of the type just described. Each D_n is a formal power series in $\hbar$, and if each $\mathcal{O}_i = O(\hbar^0)$, then it can be shown that $D_n(\otimes \mathcal{O}_i) = O(\hbar)$, essentially because there are no ambiguities of any kind in the underlying classical theory. The expression $D(\mathrm{e}_\otimes^F)$ may be viewed as being equal to the finite counterterms that characterize the difference between the two prescriptions for the time ordered products. Note that in curved space, there is even an ambiguity in defining time-ordered products with one factor (the Wick powers), so even D_1 might be non-trivial.

The counterterms, i.e., the maps D_n, satisfy a number of properties corresponding to the properties T1—T11 of the time ordered products [62]. As we have already said, the D_n are supported on the total diagonal, and this corresponds to the causal factorization property T8. The D_n are local and covariant functionals of the field ϕ, the metric, and m^2, in the following sense: Let $\psi : M \to M'$ be any causality and orientation preserving isometric embedding, i.e.,

$\psi^* g' = g$. If D_n and D'_n denote the functionals on M respectively M', then we have that $\psi^* \circ D'_n = D_n \circ (\psi^* \otimes \cdots \otimes \psi^*)$. This follows from T1. It follows from the smoothness and analyticity properties T4, T5 and the scaling property T2 that the D_n depend only polynomially on the Riemann curvature tensor, the mass parameter m^2, and the field ϕ. Since there is no ambiguity in defining the identity operator, $\mathbb{1}$, or the basic field, ϕ, we must have

$$D_1(\mathbb{1}) = D_1(\phi) = 0\,. \tag{252}$$

As a consequence of the symmetry of the time-ordered products T6, the maps D_n are symmetric (respectively graded symmetric when Grassmann valued fields would be present), and as a consequence the field independence property T9, they must satisfy

$$\frac{\delta}{\delta\phi(y)} D_n\Big(\mathcal{O}_1(x_1) \otimes \cdots \otimes \mathcal{O}_n(x_n)\Big) = \sum_k D_n\Big(\mathcal{O}_1(x_1) \otimes \cdots \frac{\delta \mathcal{O}_k(x_k)}{\delta\phi(y)} \otimes \cdots \mathcal{O}_n(x_n)\Big)\,. \tag{253}$$

In particular, the D_n depend polynomially upon the field ϕ. As a consequence of the scaling property T2 of time-ordered products, the engineering dimension of each term appearing in D_n must satisfy the following constraint. As above, let $\mathcal{N}_r$ the counter of Riemann curvature tensors, let $\mathcal{N}_f$ be the dimension counter for the fields, and let $\mathcal{N}_c$ be the counter for the coupling constant (in this case m^2), see eq. (140). Let the dimension counter $\mathcal{N}_d : \mathbf{P} \to \mathbf{P}$ be defined as above by $\mathcal{N}_d = \mathcal{N}_c + \mathcal{N}_r + \mathcal{N}_f$ Then we must have

$$(\mathcal{N}_d + sd) D_n\Big(\mathcal{O}_1(x_1) \otimes \cdots \otimes \mathcal{O}_n(x_n)\Big) = \sum_{i=1}^{n} D_n\Big(\mathcal{O}_1(x_1) \otimes \cdots \mathcal{N}_d \mathcal{O}_i(x_i) \otimes \ldots \mathcal{O}_n(x_n)\Big)\,. \tag{254}$$

where sd is the scaling degree, see appendix C. The unitarity requirement T7 on the time-ordered products yields the constraint

$$D_n\Big(\mathcal{O}_1(x_1) \otimes \cdots \otimes \mathcal{O}_n(x_n)\Big)^* = -D_n\Big(\mathcal{O}_1(x_1)^* \otimes \cdots \otimes \mathcal{O}_n(x_n)^*\Big)\,. \tag{255}$$

and the action Ward identity T11 implies that one can freely pull an exterior derivative $d_i = dx_i^\mu \wedge \frac{\partial}{\partial x_i^\mu}$ into D_n,

$$d_i D_n\Big(\mathcal{O}_1(x_1) \otimes \cdots \mathcal{O}_i(x_i) \otimes \ldots O_n(x_n)\Big) = D_n\Big(\mathcal{O}_1(x_1) \otimes \cdots d_i\, \mathcal{O}_i(x_i) \otimes \ldots O_n(x_n)\Big)\,. \tag{256}$$

The meaning of the above restrictions on D_n is maybe best illustrated in some examples. The dimension of the coupling is $d(m^2) = +2$, and the dimension of the field is $d(\phi) = +1$. Consider the composite field $\phi^2 \in \mathbf{P}$. In curved spacetime, there is an ambiguity $D_1(\phi^2)$ in defining $T_1(\phi^2)$, given by

$$\hat{T}_1(\phi^2) = T_1(\phi^2) + (\hbar/i)\, T_1(D_1(\phi^2))\,. \tag{257}$$

By properties (253) and (252), we must have $\frac{\delta}{\delta\phi}D_1(\phi^2) = 0$, so $D_1(\phi^2)$ must be a multiple of the identity operator, so $D_1(\phi^2) = ic\mathbb{1}$. By the local and covariance property and the dimensional constraint (254), $c = aR + bm^2$, where a, b are constants that must be real in view of (255). Thus, we have the familiar result that the Wick power $T_1(\phi^2)$ is unique only up to curvature/mass terms. Consider next the ambiguity in defining the time ordered product of two factors of ϕ^2, given by

$$\hat{T}_2(\phi^2 \otimes \phi^2) = T_2(\phi^2 \otimes \phi^2) + (\hbar/i)^2\, T_1(D_2(\phi^2 \otimes \phi^2)) \tag{258}$$

(here we are assuming that $D_1(\phi^2) = 0$ for simplicity). By the same reasoning as above, this must now be given by

$$D_2(\phi^2(x) \otimes \phi^2(y)) = c\delta(x, y) \tag{259}$$

for some real constant c, because the scaling degree of the delta function in 4 dimensions is $+4$. If ϕ^2 in this formula would be replaced by ϕ^3, then the right side could be a constant times the wave operator $\Box$ of the delta function, or by a real linear combination of m^2, R and ϕ^2, times the delta-function.

We summarize the renormalization ambiguities again in the "main-theorem of renormalization theory:"

Theorem 2. [62, 63] Time ordered products T with the above properties T1-T11 exist. If $T = \{T_n\}$ and $\hat{T} = \{\hat{T}_n\}$ are two different time ordered products satisfying conditions T1–T11, then their difference is given by

$$\hat{T}_n\Big(O_1(x_1) \otimes \cdots \otimes O_n(x_n)\Big) = \sum_{I_0 \cup I_1 \cup \ldots I_r \subset \underline{n}} T_{r+1}\Big(\bigotimes_{j \in I_0} O_j(x_j) \otimes \bigotimes_k (\hbar/i)^{|I_k|} D_{|I_k|}\Big[\bigotimes_{i \in I_k} O_i(x_i)\Big]\Big)\,. \tag{260}$$

Here, the sum runs over all partitions $I_0 \cup \cdots \cup I_r = \underline{n}$ of $\underline{n} = \{1, \ldots, n\}$, and $D = \{D_n\}$ is a hierarchy of counterterms described above. Conversely, if D is as above, then $\hat{T}$ defines a new hierarchy of time-ordered products with the properties T1—T11.

3.7 Perturbative construction of interacting quantum fields

In the previous sections we have given the construction of Wick powers and their time-ordered products in a theory that is classically described by a Lagrangian $\mathbf{L}_0$ at most quadratic in the field, with associated classical field equations of wave-equation type. Those quantities may be used to give a definition of an interacting quantum field theory via a perturbation expansion. For definiteness, consider a scalar field described by the classical Lagrangian $\mathbf{L} = \mathbf{L}_0 + \lambda\mathbf{L}_1$,

$$\mathbf{L} = \frac{1}{2}(d\phi \wedge *d\phi + m^2 *\phi^2) + \lambda *\phi^N = \mathbf{L}_0 + \lambda\mathbf{L}_1\,. \tag{261}$$

We would like to construct quantities in the interacting quantum field theory as formal power series in λ. Even in flat spacetime, one may encouter infra-red divergences if one tries to define

the terms in such expansions, but such infra-red divergences are absent if one considers, instead of the interaction $I = \int \lambda \mathbf{L}_1$, a cutoff interaction, $F = \int \lambda f \mathbf{L}_1$, where f is a smooth cutoff function of compact support that is one in a globally hyperbolic subregion of the original spacetime (M,g). The perturbative formula for the interacting fields associated with this interaction is then

$$\mathcal{O}(x)_F = T\left(\mathrm{e}_{\otimes}^{iF/\hbar}\right)^{-1} \star_{\hbar} \frac{\delta}{\delta j(x)} T\left(\mathrm{e}_{\otimes}^{iF/\hbar + \int j \wedge \mathcal{O}}\right)\Big|_{j=0}. \tag{262}$$

This formula is called "Bogoliubov's formula," [12]. Each term in the formal power series for $\mathcal{O}(x)_F$ is a well-defined element in $\mathbf{W}_0$, due to the infra-red cutoff in the interaction F. The subscript "F" indicates throughout this paper the we mean an "interacting field" defined by F, which is an element in the ring[10] $\mathbf{W}_0 \otimes \mathbb{C}[[\lambda,\hbar]]$, as opposed to the classical field expression $\mathcal{O} \in \mathbf{P}$. The expansion coefficients in λ of the interacting fields define the so-called "retarded products," [79]

$$\mathcal{O}(x)_F = \sum_{n=0}^{\infty} \frac{i^n}{\hbar^n n!} R_n(\mathcal{O}(x); F^{\otimes n}) =: R\big(\mathcal{O}(x); \mathrm{e}_{\otimes}^{iF/\hbar}\big). \tag{263}$$

The retarded products are maps $R_n : \mathbf{P}^{\otimes(n+1)} \to \mathcal{D}'(M^{n+1}) \otimes \mathbf{W}_0$ with properties similar to the properties T1—T11 of the time-ordered products. The symmetry property only holds with respect to the n-arguments separated by the semicolon. Their definition in terms of time-ordered products is

$$\begin{aligned} & R_n\Big(\Psi(y); \mathcal{O}_1(x_1) \otimes \cdots \otimes \mathcal{O}_n(x_n)\Big) \\ = & \sum_{I_1 \cup \cdots \cup I_j = \underline{n}} (-1)^{n+j+1} T_{|I_1|}\Big(\bigotimes_{k \in I_1} \mathcal{O}_k(x_k)\Big) \star_{\hbar} \ldots \star_{\hbar} T_{|I_j|}\Big(\Psi(y) \otimes \bigotimes_{k \in I_j} \mathcal{O}_k(x_k)\Big), \end{aligned} \tag{264}$$

where the sum runs over all partitions $I_1 \cup \cdots \cup I_j$ of $\underline{n} = \{1,\ldots,n\}$. An important property of the retarded product is that their support is restricted to the set

$$\operatorname{supp} R_n(\Psi(y); \mathcal{O}_1(x_1) \otimes \cdots \otimes \mathcal{O}_n(x_n)) \subset \{(y, x_1, \ldots, x_n) \in M^{n+1} \mid x_i \in J^-(y) \quad \forall i\}. \tag{265}$$

The support property follows from the causal factorization property of the time-ordered products. A useful combinatorial identity for the retarded products is the Glaser-Lehmann-Zimmermann (GLZ) relation, which states that [35]

$$\begin{aligned} & R_n\Big(\Psi_1(y_1); \Psi_2(y_2) \otimes \bigotimes_{i=1}^{n-1} \mathcal{O}_i(x_i)\Big) - R_n\Big(\Psi_2(y_2); \Psi_1(y_1) \otimes \bigotimes_{i=1}^{n-1} \mathcal{O}_i(x_i)\Big) = \\ & \sum_{I \cup J = \underline{n}} \Big[R_{|I|}\Big(\Psi_1(y_1); \bigotimes_{i \in I} \mathcal{O}_i(x_i)\Big), R_{|J|}\Big(\Psi_2(y_2); \bigotimes_{j \in J} \mathcal{O}_j(x_j)\Big)\Big] \end{aligned} \tag{266}$$

[10] The fact that, implicit in the notation "$\mathbb{C}[[\hbar]]$", the interacting field only contains non-negative powers of $\hbar$, is not so obvious and follows from the fact that R_n itself is of order $\hbar^n$, see [38].

The GLZ-relation may be used to express the commutator of two interacting fields in terms of retarded products as follows:

$$[\Psi_1(y_1)_F, \Psi_2(y_2)_F] = \sum_{n=0}^{\infty} \frac{i^n}{\hbar^n n!} \Big[R_{n+1}(\Psi_1(x_1); \Psi_2(x_2) \otimes F^{\otimes n}) - (1 \leftrightarrow 2) \Big] . \tag{267}$$

As a consequence of the GLZ-relation and the support properties of the retarded products, any two interacting fields located at spacelike separated points commute[11]. Thus, we have constructed interacting fields as formal power series in the coupling constant via the time-ordered products in the underlying free field theory. If one changes the definition of the time-ordered products along the lines described in the previous subsection, then there is a corresponding change in the interacting theory, affecting both the interaction Lagrangian, as well as resulting in general in a multiplicative redefinition of the interacting fields. To describe this in more detail, we introduce the linear map $\mathbf{Z}_F : \mathbf{P}(M) \to \mathbf{P}(M)[[\lambda, \hbar]]$ by

$$\mathbf{Z}_F(\mathcal{O}(x)) := \mathcal{O}(x) + D(\mathcal{O}(x) \otimes \mathrm{e}_{\otimes}^{F}) , \tag{268}$$

where $D = \{D_n\}$ is the hierarchy of distributions encoding the difference between two prescriptions T and $\hat{T}$ for time ordered products. We may introduce a basis in $\mathbf{P}(M)$, and represent this map by its matrix

$$\mathbf{Z}_F(\mathcal{O}_i(x)) = \sum_j Z_i^j \, \mathcal{O}_j(x) . \tag{269}$$

For renormalizable interactions ($\mathcal{N}_f F \leq 4$), $\mathbf{Z}_F$ leaves each finite dimensional subspace of $\mathbf{P}$ invariant, but this is no longer the case for non-renormalizable interactions. Now, if $\hat{O}(x)_F$ is the definition of the interacting field using the time ordered products $\hat{T}$, and $\mathcal{O}(x)_F$ that using T, then the two are related by

$$\hat{\mathcal{O}}(x)_F = \mathbf{Z}_F[\mathcal{O}(x)]_{F + D(\exp_{\otimes} F)} . \tag{270}$$

We now explain how one can remove the cutoff implemented by the cutoff function f in the interaction $F = \int \lambda f \mathcal{O}$ at the algebraic level. The key identity [18] in this construction is

$$V_{F_1,F_2} \star_\hbar \mathcal{O}(x)_{F_2} \star_\hbar V_{F_1,F_2}{}^{-1} = \mathcal{O}(x)_{F_1} \tag{271}$$

where F_1, F_2 are any two local interactions as above that are equal in an open neighborhood of x, and where $V_{F_1,F_2} \in \mathbf{W}_0 \otimes \mathbb{C}[[\hbar, \lambda]]$ are unitaries that can be written in terms of retarded products. They satisfy the cocycle condition

$$V_{F_1,F_2} \star_\hbar V_{F_2,F_3} \star_\hbar V_{F_3,F_1} = 1\!\!1 . \tag{272}$$

[11] In case when Grassmann valued fields are present, the commutator is replaced by the graded commutator, and the minus sign on the right side is replaced by $-(-1)^{\varepsilon_1 \varepsilon_2}$, where ε_i are the Grassmann parities of Ψ_i.

To construct the limit of the interacting fields as $f \to 1$, one can now proceed as follows. For simplicity, let us assume that $M = \mathbb{R} \times \Sigma$, with Σ compact. The cutoff function may then be chosen to be of compact support in a "time-slice" $M_{2\tau} = \Sigma \times (-2\tau, 2\tau)$, and to be equal to one in a somewhat smaller time-slice, say M_τ. To indicate the dependence upon the cutoff τ, let us write the cutoff function as f_τ, and let us correspondingly write. Let $F_\tau = \int \lambda f_\tau \mathbf{L}_1$ and $\mathcal{O}_{F_\tau}$ for the corresponding interacting field defined using F_τ as the interaction. Finally, let $U_\tau = V_{F_\delta, F_\tau}$, for some fixed δ. The interacting fields defined with respect to the true interaction $I = \int \lambda \mathbf{L}_1$ may now defined as the limit

$$\mathcal{O}(x)_I = \lim_{T \to \infty} U_\tau \star_\hbar \mathcal{O}(x)_{F_\tau} \star_\hbar {U_\tau}^{-1} . \tag{273}$$

The sequence on the right side is trivially convergent, because it only contains a finite number of terms for each fixed x, by the cocycle condition. More precisely, the terms in the sequence will remain constants once τ has become so large that $x \in M_\tau$. It is important to note that this would not be the case if we had not inserted the unitary operators under the limit sign. In that case, our notion of interacting field would have coincided with the naive "adiabatic limit" which intuitively corresponds to the situation where the interacting field is fixed at $\tau = -\infty$. By contrast, our limit corresponds intuitively to fixing the field during"finite time interval" corresponding to the neighborhood $\Sigma \times (-\delta, \delta)$. Actually, one can see that the defining formula for U_τ and the interacting field will still make sense also for spacetimes with non-compact Cauchy surface. We can now define the algebras of interacting field observables as

$$\mathcal{F}_I(M, g) = \mathrm{Alg}\left\{ G_I \;\middle|\; G = \int g \wedge \mathcal{O} \right\} \Big/ \mathcal{J}_0 . \tag{274}$$

We note that these are subalgebras of $\mathcal{F}_0[[\lambda, \hbar]]$. While the embedding of this algebra as a subalgebra of $\mathcal{F}_0[[\lambda, \hbar]]$ depends upon the choice of the cutoff function f, it can be proved [18, 64] that the definition of $\mathcal{F}_I$ as an abstract algebra is independent of our choice of the sequence of cutoff functions $\{f_\tau\}$. Another important consequence of our definition of the interacting fields is that, if we want to investigate properties of the interacting field near a point x, we only have to work in practice with the cutoff interaction F where f is equal to 1 on a sufficiently large neighborhood containing x. For example, if we want to check whether an interacting current $\mathbf{J}(x)_I$ is conserved, we only need to check whether $d\mathbf{J}(x)_F = 0$ for any cutoff function f which is equal to 1 in an open neighborhood of x.

The effect of changing the renormalization conditions may also be discussed at the level of the interacting fields $\mathcal{O}_I$ and the associated interacting field algebra $\mathcal{F}_I$. For this, consider again two prescriptions T and $\hat{T}$ for defining the time-ordered products, and let us denote by $\mathcal{O}_I$ and $\hat{O}_I$ the respective interacting fields, and by $\mathcal{F}_I$ and $\hat{\mathcal{F}}_I$ the interacting field algebras. Let us denote by $\mathbf{Z}_I : \mathbf{P} \to \mathbf{P}[[\lambda, \hbar]]$ the limit of the map $\mathbf{Z}_F$ as the cutoff implicit in F is removed. This limit exists, because all the functionals $D = \{D_n\}$ in the defining relation (268) for $\mathbf{Z}_F$ are supported only on the total diagonal. Then one can derive from eq. (270) that there exists an algebra isomorphism

$$\rho : \hat{\mathcal{F}}_I \to \mathcal{F}_{\hat{I}}, \quad \rho(\hat{\mathcal{O}}_I) = \mathbf{Z}_I(\mathcal{O})_{\hat{I}} , \tag{275}$$

with $\hat{I} = I + D(e^I_\otimes)$. The algebra isomorphism map ρ is needed in order to compensate for the difference between the unitaries U_τ and $\hat{U}_\tau$ in the two prescriptions, see eq. (273), and see [64] for details. A particular case of this map again arises when the prescription $\hat{T}$ is defined in terms of a change of scale (see T2) from the time ordered product T. Then we obtain, for each scale $\mu \in \mathbb{R}^+$, a map ρ_μ, which depends polynomially on μ and $\ln \mu$. This map defines the renormalization group flow in curved spacetime [64] together with the corresponding "mixing matrices," i.e., the matrix components $Z^i_j(\mu)$ of the maps $Z_I(\mu)$.

4 Quantum Yang-Mills theory

4.1 General outline of construction

4.1.1 Free fields

We now construct quantum Yang-Mills theory along the lines outlined in the introduction. As our starting point, we take the auxiliary theory described classically by the auxiliary action S with ghosts and anti-fields, see eq. (34). Thus, the set of dynamical and background fields is

background fields	dynamical fields
spacetime metric g	
anti-ghost $C^\ddagger, \bar{C}^\ddagger$	ghost $C, \bar{C}$
anti-vector $A^\ddagger$	vector A
anti-auxiliary $B^\ddagger$	auxiliary B

We assume that the group G is a direct product of a semi-simple group and $U(1)^l$, and that the dimension of spacetime is 4. We split the action S into a free part S_0 containing only expressions at most quadratic in the dynamical fields, and an interaction part, $\lambda S_1 + \lambda^2 S_2$. The action S_0 describes the classical auxiliary theory. Its field equations are hyperbolic. As we shall describe in more detail below, we can thus define an algebra $\mathbf{W}_0$ that represents a deformation quantization of the free field theory associated with the free auxiliary action S_0, and this algebra contains all local covariant Wick-powers, and their time-ordered products.

As in the classical case, the so-obtained auxiliary theory is by itself not equivalent to (free) Yang-Mills theory, because it contains gauge-variant observables and observables with non-zero ghost number. To obtain a quantum theory of (free) Yang-Mills theory, we pass from the algebra of observables, $\mathbf{W}_0$, to the cohomology algebra constructed from the (free) quantum BRST-charge Q_0. For this, we consider first the (free) classical BRST-current $\mathbf{J}_0$, which defines a quantum Wick power $T_1(\mathbf{J}_0)$, which we denote again by J_0 by abuse of notation. Let us assume for simplicity that the spacetime (M,g) has a compact Cauchy surface Σ. Then there is a closed compactly supported 1-form γ on M such that $\int_M \gamma \wedge \alpha = \int_\Sigma \alpha$ for any closed 3-form α, i.e., $[\gamma] \in H^1_0(M,d)$ is dual to the cycle $[\Sigma] \in H_3(M,\partial)$. We can then define the *free* BRST-charge by

$$Q_0 = \int_M \gamma \wedge \mathbf{J}_0 \tag{276}$$

As we will show below, the local covariant quantum BRST current $\mathbf{J}_0 := T_1(\mathbf{J}_0)$ can be defined so that it is closed $d\mathbf{J}_0 = 0$ modulo $\mathcal{J}_0$, so evidently Q_0 is independent, modulo $\mathcal{J}_0$, of the choice of the representer γ in $H^1(M,d)$. We will also show that Q_0 is nilpotent, $Q_0^2 = 0$ modulo $\mathcal{J}_0$. It follows from this fact that the linear quotient space

$$\hat{\mathcal{F}}_0 = \frac{\text{Kernel}\,[Q_0, \,.\,] \cap \mathcal{F}_0 \cap \text{Kernel}\,\mathcal{N}_g}{\text{Image}\,[Q_0, \,.\,] \cap \mathcal{F}_0 \cap \text{Kernel}\,\mathcal{N}_g}\,, \quad \mathcal{F}_0 = \mathbf{W}_0/\mathcal{J}_0 \tag{277}$$

is well defined, and that it is again an algebra. Above, we have explained that $\mathcal{F}_0$ is a deformation quantization of the classical theory associated with S_0 in the sense that, when $\hbar \to 0$, the commutator divided by $\hbar$ goes over to the Peierls bracket of the classical observables. In particular, the commutator divided by $\hbar$ with Q_0 goes to the classical BRST-variation, $\hat{s}_0$. Furthermore, as we explained above, the cohomology of $\hat{s}_0$ is in 1-1 correspondence with classical gauge-invariant observables, so that, in the classical limit, the algebra $\hat{\mathcal{F}}_0$ is the Poisson algebra of physical, gauge-invariant observables. Thus, it is natural to define $\hat{\mathcal{F}}_0$ to be the algebra of physical observables also in the quantum case.

Consider now a representation π_0 of the free algebra $\mathcal{F}_0$ on an inner product space $\mathcal{H}_0$. For simplicity, let us denote representer $\pi_0(Q_0)$ of the BRST-charge in this representation again by Q_0. We require Q_0 to be hermitian with respect to the (necessarily indefinite) inner product. We would like to know under which condition this representation induces a Hilbert-space representation $\hat{\pi}_0$ on the factor algebra $\hat{\mathcal{F}}_0$. Following [39], let us suppose that the representation fulfills the following additional

Positivity requirement: A representation is called positive if the following hold: (a) if $|\psi\rangle \in \text{Kernel}\,Q_0$, then $\langle\psi|\psi\rangle \geq 0$, and (b) if $|\psi\rangle \in \text{Kernel}\,Q_0$, then $\langle\psi|\psi\rangle = 0$ if and only if $|\psi\rangle \in \text{Image}\,Q_0$.

It is elementary to see that if the positivity requirement is fulfilled, then the representation π_0 induces a representation $\hat{\pi}_0$ of the physical observables $\hat{\mathcal{F}}_0$ on the inner product space

$$\hat{\mathcal{H}}_0 = \frac{\text{Kernel}\,Q_0}{\text{Image}\,Q_0}\,, \tag{278}$$

which is in fact seen to be a pre-Hilbert space, i.e., carries a positive definite inner product. As we will see below, when G is compact, there do indeed exist representations satisfying the above positivity requirement if we restrict ourselves to the ghost number 0 subalgebra of $\mathcal{F}_0$. As we will also see, in static spacetimes (M,g) or in spacetimes with static regions, the states in $\hat{\mathcal{H}}_0$ (in the ground state representation) can be put into one-to-one correspondence with $\pm$-helicity particle states of the electromagnetic field, and $\hat{\mathcal{H}}_0$ contains a dense set of Hadamard states. However, in generic time-dependent spacetimes, such an interpretation in terms of particles states is not possible.

When the Cauchy surfaces of M are not compact, the charge Q_0 is in general not defined as stated. The reason is that the 1-form field γ is no longer of compact support, but has non-compact support in spatial directions. Nevertheless, we can see that if we formally consider the graded commutator $[Q_0, \mathcal{O}(x)]$ with a local quantum Wick-power, denoted $\mathcal{O}(x) := T_1(\mathcal{O}(x))$, then there will be only contributions in the formal integral defining Q_0 (see (276)) from the portion of the support of γ that is contained in $J^+(x) \cup J^-(x)$. All other contributions vanish due to the (graded) commutativity property, T9. Since the intersection of the support of γ and $J^+(x) \cup J^-(x)$ is compact for a suitable choice of γ, it follows that the commutator of any local observable in $\mathcal{F}_0$ with Q_0 is always defined. Thus, while Q_0 itself is undefined, the graded commutator still defines a graded derivation. The definition of the algebra of gauge invariant observables can then be given in terms of this graded derivation. However, the construction of representations explicitly used (the representer of) Q_0 itself, and not just the graded commutator. Thus, it is not straightforward to obtain Hilbert space representations on manifolds with non-compact Cauchy surfaces.

4.1.2 Interacting fields

A similar kind of construction as for free Yang-Mills theory can also be given in order to perturbatively construct quantized interacting Yang-Mills theory. The starting point is now the classical auxiliary interacting field theory described by the auxiliary action $S = S_0 + \lambda S_1 + \lambda^2 S_2$. Thus, the interaction is

$$I = \int (\lambda \mathbf{L}_1 + \lambda^2 \mathbf{L}_2) = \lambda S_1 + \lambda^2 S_2 \,. \tag{279}$$

The first step is to construct a quantum theory associated with this auxiliary action. For simplicity, we again assume that M has compact Cauchy-surfaces—the general situation can again be treated by complete analogy with the free field case as just described. Following the general procedure described in Sec. 3.7, we first introduce an infra-red cutoff for the interaction supported in a compact region of spacetime, and construct the interacting theory in that region. To define the desired infra-red cutoff, we consider a compactly supported cutoff function, f, which is equal to 1 on the submanifold $M_\tau = (-\tau, \tau) \times \Sigma$. We define a cutoff interaction, F, by $F = \int \{ f \lambda \mathbf{L}_1 + f^2 \lambda^2 \mathbf{L}_2 \}$, and we define corresponding interacting fields $\mathcal{O}_F$ by Bogoliubov's formula. We then send the cutoff τ to infinity at the algebraic level as described in sec. 3.7, and get a corresponding algebra $\mathcal{F}_I$ of interacting fields $\mathcal{O}_I$. This algebra of interacting fields is not equivalent to quantum Yang-Mills theory, as it contains gauge variant fields and fields of non-zero ghost number. As in the free case, we obtain the algebra of physical field observables by considering the cohomology of the (now interacting) BRST-charge operator, Q_I.

To define this object, consider the interacting BRST-current with cutoff interaction, defined

by the Bogoliubov formula [see eq. (263)]

$$\begin{aligned}\mathbf{J}(x)_F &= \frac{\delta}{\delta\gamma(x)} T(e_\otimes^{iF/\hbar})^{-1} \star_\hbar T(e_\otimes^{iF/\hbar+\int\gamma\wedge\mathbf{J}})\Big|_{\gamma=0} \\ &= \sum_{n\geq 0}\frac{1}{n!}\left(\frac{i}{\hbar}\right)^n R_n(\mathbf{J}(x);F^{\otimes n})\,. \end{aligned} \tag{280}$$

As in our general definition of interacting fields, we can then remove the cutoff at the algebraic level by defining an interacting current $\mathbf{J}(x)_I$. We will show below that the interacting BRST-current $\mathbf{J}_I(x)$ is conserved in M, so we can define a corresponding interacting BRST-charge by $Q_I = \int \gamma\wedge\mathbf{J}_I$, [compare eq. (276)].

We will furthermore show that the so-defined charge is nil-potent, $Q_I^2 = 0$. Thus, we can define the physical observables as in the free field theory by the cohomology of the interacting BRST-charge, i.e., the algebras of interacting fields are defined by

$$\hat{\mathcal{F}}_I = \frac{\mathrm{Kernel}\,[Q_I,\,.\,]\cap\mathcal{F}_I\cap\mathrm{Kernel}\,\mathcal{N}_g}{\mathrm{Image}\,[Q_I,\,.\,]\cap\mathcal{F}_I\cap\mathrm{Kernel}\,\mathcal{N}_g}\,. \tag{281}$$

Next, one would like to define representations of the algebra of observables on a Hilbert space. Such representations can be obtained from those of the free theory by a deformation process [39]. For this, consider a state $|\psi_0\rangle \in \mathcal{H}_0$ in a representation π_0 of the underlying free theory satisfying the above positivity requirement. Let also $|\psi_0\rangle \in \mathrm{Kernel}\,Q_0$. Then, using $Q_I^2 = 0$, and $Q_I = Q_0 + \lambda Q_1 + \lambda^2 Q_2 + \dots$ one first shows that there exists a formal power series

$$|\psi_I\rangle = |\psi_0\rangle + \lambda|\psi_1\rangle + \lambda^2|\psi_2\rangle + \dots \in \mathcal{H}_I = \mathcal{H}_0[[\lambda]] \tag{282}$$

such that $Q_I|\psi_I\rangle = 0$, where Q_I has been identified with its representer in the representation π_I that is induced from the representation of the underlying free theory. In order to construct the vectors $|\psi_i\rangle$, we proceed inductively. We write the condition that $|\psi_I\rangle$ is in the kernel of Q_I and that $Q_I^2 = 0$ as

$$0 = \sum_{k=0}^{m} Q_k|\psi_{m-k}\rangle\,, \quad 0 = \sum_{k=0}^{m} Q_k Q_{m-k}\,, \tag{283}$$

for all m. For $m = 0$, the first equation is certainly satisfied, as we are assuming $Q_0|\psi_0\rangle = 0$. Assume now that $|\psi_0\rangle, |\psi_1\rangle, \dots, |\psi_{n-1}\rangle$ have been constructed in such a way that the first equation is satisfied up to $m = n-1$, and put

$$|\chi_m\rangle = \sum_{k=0}^{n-1} Q_{m-k}|\psi_k\rangle\,. \tag{284}$$

Then, using the second equation in (283), we see that

$$0 = \sum_{k=0}^{m} Q_{m-k}|\chi_k\rangle\,, \quad 0 = \sum_{k=0}^{m} \langle\chi_m|\chi_{m-k}\rangle\,, \tag{285}$$

for all m. We now use the inductive assumption that $|\chi_m\rangle = 0$ for $m \le n-1$, from which we get that $Q_0|\chi_n\rangle = 0$, putting $m = n$ in the first equation. Putting $m = 2n$ in the second equation, we get $\langle\chi_n|\chi_n\rangle = 0$. In view of the positivity requirement, we must thus have $|\chi_n\rangle = -Q_0|\psi_n\rangle$ for some $|\psi_n\rangle$. We take this as the definition of the n-th term for the deformed state (282). This then satisfies the induction assumption at order n, thus closing the induction loop.

Thus, by the above deformation argument, one sees that $\mathrm{Kernel}\, Q_I \subset \mathcal{H}_I$ is a non-empty subspace. One furthermore shows that the representation π_I satisfies an analog of the positivity requirement[12] for the interacting theory. Thus, we obtain, as in the free case, a representation $\hat{\pi}_I$ on the inner product space

$$\hat{\mathcal{H}}_I = \frac{\mathrm{Kernel}\, Q_I}{\mathrm{Image}\, Q_I}\,, \tag{286}$$

and this space is again shown to be a pre-Hilbert space. For details of these constructions, see sec. 4.3 of [39].

4.1.3 Operator product expansions and RG-flow

As we have just described, a physical gauge invariant, interacting field is an element in the algebra $\hat{\mathcal{F}}_0$, i.e., an equivalence class of an interacting field operator $\mathcal{O}_I(x)$ satisfying

$$[Q_I, \mathcal{O}_I(x)] = 0 \quad \forall x \in M\,, \tag{287}$$

modulo the interacting fields that can be written as

$$\mathcal{O}_I(x) = [Q_I, \mathcal{O}'_I(x)] \quad \forall x \in M\,, \tag{288}$$

for some local field $\mathcal{O}'$ (as usual, $[\,,\,]$ means the graded commutator). Our constructions of the interacting BRST-charge do not imply that the action of Q_I on a local covariant interacting field is not equivalent to $\hat{s}$. But it follows from general arguments that

$$[Q_I, \mathcal{O}_I(x)] = (\hat{q}\mathcal{O})_I(x) \quad \forall x \in M \tag{289}$$

where $\hat{q}$ is a map

$$\hat{q} : \mathbf{P}^p(M) \to \mathbf{P}^p(M)[[\hbar]], \quad \hat{q} = \hat{s} + \hbar\hat{q}_1 + \hbar^2\hat{q}_2 + \dots\,. \tag{290}$$

Because $Q_I^2 = 0$, the map $\hat{q}$ is again a differential (the "quantum BRST-differential"), $\hat{q}^2 = 0$, whose action on general elements in $\mathbf{P}$ is different from that of $\hat{s}$. An exception of this rule are the exactly gauge invariant elements $\mathcal{O} = \mathbf{\Psi}$ at zero ghost number, which by lemma 1 are of the

[12] Since we are working over the ring $\mathbb{C}[[\lambda]]$ of formal power series in λ in the case of interacting Yang-Mills theory, the positivity requirement needs to be formulated appropriately by specifying what it means for a formal power series to be positive. For details, see [39].

form $\Psi = \prod \Theta_{s_i}(F, \mathcal{D}F, \mathcal{D}^2 F, \dots)$, with Θ_s invariant polynomials of the Lie-algebra. For such elements, we shall show that we have $\hat{q}\Psi = \hat{s}\Psi = 0$. Thus,

$$[Q_I, \Psi_I(x)] = 0 \quad \forall x \in M \tag{291}$$

and the corresponding interacting fields $\Psi_I(x)$ are always observable.

Given n local fields $\mathcal{O}_{j_1}, \dots, \mathcal{O}_{j_n} \in \mathbf{P}$, we can construct the operator product expansion of the corresponding interacting quantum fields,

$$\mathcal{O}_{j_1}(x_1)_I \star_\hbar \dots \star_\hbar \mathcal{O}_{j_n}(x_n)_I \sim \sum_k C^k_{j_1 \dots j_n}(x_1, \dots, x_n, y)\, \mathcal{O}_k(y)_I\,. \tag{292}$$

The operator product expansion is an asymptotic expansion for $x_1, \dots, x_n \to y$, see [68], where the construction and properties of the expansion are described. Because the action S of the auxiliary theory has zero ghost number, the OPE coefficients are non-vanishing only when

$$\sum_r \mathcal{N}_g(\mathcal{O}_{j_r}) = \mathcal{N}_g(\mathcal{O}_k)\,. \tag{293}$$

Now assume that all operators $\mathcal{O}_{j_1}, \dots, \mathcal{O}_{j_n}$ are physically observable fields. Then, since the graded commutator with Q_I respects the $\star_\hbar$-product, also all local operators $\mathcal{O}_k$ appearing on the right side must be in the kernel of Q_I. By the same argument, if one of the operators on the left side is of the trivial from (288), then it follows that each operator in the expansion on the right side is of that form, too. Thus, we conclude that the OPE closes on gauge invariant operators, and we summarize this important result as a theorem:

Theorem 3. Let $\mathcal{O}_{i_1}, \dots, \mathcal{O}_{i_n} \in \mathbf{P}$ be in the kernel of $\hat{s}$, with vanishing ghost number, as characterized by thm. 1. Then $C^k_{i_1 \dots i_n}$ is non-vanishing only for $\mathcal{O}_k \in \mathbf{P}$ of vanishing ghost number that are in the kernel of $\hat{s}$. If one $\mathcal{O}_{i_r}$ is in the image of $\hat{s}$, then $C^k_{i_1 \dots i_n}$ is non-vanishing only for $\mathcal{O}_k \in \mathbf{P}$ of vanishing ghost number that are in the image of $\hat{s}$. If one drops the restriction to the 0-ghost number sector, then the same statement is true with $\hat{s}$ replaced by $\hat{q}$.

By the same kind of argument, one can also show that the renormalization group flow closes on physical operators. The renormalization flow in curved spacetime was defined in subsec. 3.7 as the behavior of the interacting fields under a conformal change of the metric, $g \to \mu^2 g$. In general we have $\rho_\mu(\mathcal{O}_i(x)_I(x)) = Z_i^j(\mu) \cdot \mathcal{O}_j(x)_{I_\mu}$ for all $x \in M$, where I_μ is the renormalized interaction, and where $\rho_\mu : \mathcal{F}_I(g) \to \mathcal{F}_{I_\mu}(\mu^2 g)$ is an algebraic isomorphism implementing the conformal change of the metric. Now, in the perturbative quantum field theory associated with the auxiliary action S, we have

$$\rho_\mu(\mathbf{J}(x)_I) = Z(\mu) \cdot \mathbf{J}(x)_{I_\mu} + \sum_i \zeta_i(\mu) \cdot \mathcal{O}_i(x)_{I_\mu} \quad \forall x \in M\,, \tag{294}$$

for some $Z(\mu), \zeta_i(\mu) \in \mathbb{C}[[\lambda, \hbar]]$, and operators $\mathcal{O}_i \in \mathbf{P}^3(M)$ of dimension three not equal to the BRST-current and not equal to 0. If we take the exterior derivative d of this equation and use

that the interacting BRST-currents themselves are conserved, we obtain $\sum \zeta_i(\mu) \cdot d\mathcal{O}_i(x)_{I_\mu} = 0$. Let k be the largest natural number such that $\zeta_i(\mu)$ is of order $\hbar^k$ for all i, and let $z_i(\mu)$ be the $\hbar^k$-contribution to $\zeta_i(\mu)$. We can then divide this relation by $\hbar^k$, and take the classical limit $\hbar \to 0$. Because the classical limit of the interacting fields gives the corresponding perturbatively defined classical interacting fields and because $I_\mu \to I$ as $\hbar \to 0$, it follows that $\sum z_i(\mu) \cdot d\mathcal{O}_i(x)_I = 0$ for the corresponding on-shell classical interacting fields. This means that $d\mathcal{O}_i(x)_I = 0$ for those i such that $z_i(\mu) \neq 0$. But there are no such 3-form fields of dimension three at the classical level by the results of [5] except for the zero field and the BRST-current. Thus, we have found that $z_i(\mu) = 0$ for all i. By repeating this type of argument for the higher orders in $\hbar$ in $\zeta_i(\mu)$, we can conclude that $\zeta_i(\mu) = 0$ to all orders in $\hbar$.

Thus, we have found that BRST-current does not mix with other operators under the renormalization group flow, from which it follows that

$$\rho_\mu(Q_I) = Z(\mu) \cdot Q_{I_\mu}. \tag{295}$$

Hence, if $[Q_I, \mathcal{O}_i(x)_I] = 0$ for all $x \in M$, then, by applying ρ_μ to this relation, it also follows that

$$Z^i_j(\mu)[Q_{I_\mu}, \mathcal{O}_i(x)_{I_\mu}] = 0. \tag{296}$$

Because $Z^i_j(\mu)$ is invertible (it is a formal power series in λ starting with δ^i_j), we thus obtain the following result, which states that the RG-flow does not leave the sector of physical observables:

Theorem 4. Let $\mathcal{O}_i \in \mathbf{P}$ be in the kernel of $\hat{s}$, with vanishing ghost number, as characterized by thm. 1. Then $Z_i^j(\mu)$ is non-vanishing only for $\mathcal{O}_j \in \mathbf{P}$ of vanishing ghost number that are in the kernel of $\hat{s}$. If $\mathcal{O}_i$ is in the image of $\hat{s}$, then $Z_i^j(\mu)$ is non-vanishing only for $\mathcal{O}_j \in \mathbf{P}$ of vanishing ghost number that are in the image of $\hat{s}$. If one drops the restriction to the 0-ghost number sector, then the same statement is true with $\hat{s}$ replaced by $\hat{q}$.

Remark : An interesting corollary to this theorem arises when one considers the particular case when $\mathcal{O}$ is the Yang-Mills Lagrangian. Since it is the only gauge invariant field at ghost number 0 of this dimension, it does not mix with other field up to Q_I-exact terms under the renormalization group flow. The corresponding *constant* $Z_I(\mu)$ describing the field renormalization for the interacting field corresponding to the Yang-Mills Lagrangian then defines the flow of the coupling constant λ. Since our flow is local and covariant, it follows that this flow automatically must be exactly the same as in Minkowski spacetime!

A similar remark would apply to more complicated gauge theories with additional matter fields, as long as there cannot arise any additional couplings to curvature of engineering dimension 4 (such as e.g. $R \operatorname{Tr} \Phi^2$ if the gauge field is coupled to a scalar field Φ in some representation of the gauge group). Even if there can arise such couplings, the above argument can still be used to directly infer the vanishing of all β-functions in curved spacetimes with $R = 0$ if the corresponding β-functions vanish in flat spacetime.

4.2 Free gauge theory

We now describe in more detail the construction of free gauge theory outlined in the previous section 4.1. As explained, our starting point is the auxiliary theory that is classically described by the free action S_0. The first step is to define a suitable deformation quantization algebra $\mathbf{W}_0$ for this theory. The theory contains the dynamical fields $\Phi = (A^I, B^I, C^I, \bar{C}^I)$, as well as the background fields $\Phi^\ddagger = (A_I^\ddagger, B_I^\ddagger, C_I^\ddagger, \bar{C}_I^\ddagger)$. Of the dynamical fields, B^I is only an auxiliary field with no kinetic term in $S = 0$, while the vector field A^I and the ghost fields $C^I, \bar{C}^I$ were quantized above in sect. 3.5. Thus, the desired $\mathbf{W}_0$ will essentially be a tensor product of the algebras for the vector and ghost fields. We now describe the construction in detail.

We first consider a vector Hadamard 2-point function $\omega^{\mathrm{v}}(x,y)$, and a scalar Hadamard 2-point function $\omega^{\mathrm{s}}(x,y)$. These quantities by definition satisfy the hyperbolic equations

$$(d\delta + \delta d)_x \omega^{\mathrm{v}}(x,y) = 0 = (d\delta + \delta d)_y \omega^{\mathrm{v}}(x,y) \quad (d\delta)_x \omega^{\mathrm{s}}(x,y) = 0 = (d\delta)_y \omega^{\mathrm{s}}(x,y), \tag{297}$$

the commutator property (116), and the wave front condition (119). Below, we will show that we can always choose them so that they additionally satisfy the consistency relation

$$d_x \omega^{\mathrm{s}}(x,y) = -\delta_y \omega^{\mathrm{v}}(x,y), \quad d_y \omega^{\mathrm{s}}(x,y) = -\delta_x \omega^{\mathrm{v}}(x,y), \tag{298}$$

where $d_x = dx^\mu \wedge \frac{\partial}{\partial x^\mu}$, and where $\delta_x = * d_x *$ is the co-differential, etc. We define the desired deformation quantization algebra $\mathbf{W}_0$ to be the vector space generated by formal expression of the form

$$F(u) = \int u^{k_1 \dots k_n}_{i_1 \dots i_m}(x_1, \dots, x_n; y_1, \dots, y_m) :\Phi^{i_1}(y_1) \dots \Phi^{i_m}(y_m) \Phi^\ddagger_{k_1}(x_1) \dots \Phi^\ddagger_{k_n}(x_n):_\omega, \tag{299}$$

where u is a distribution subject to the wave front set condition (126) in the variables $y_1, \dots, y_m$, but not subject to any wave front set condition in the variables $x_1, \dots, x_n$. We define the $\star_\hbar$-product to be given by the differential operator

$$<\mathcal{D}> = \int \frac{\delta_L}{\delta \Phi_k(x)} \omega_{jk}(x,y) \frac{\delta_R}{\delta \Phi_j(y)} \, dx dy \tag{300}$$

where $j, k = (A^I, B^I, C^I, \bar{C}^I)$, and where

$$\big(\omega_{jk}(x,y)\big) = (k_{IJ}) \otimes \begin{pmatrix} \omega^{\mathrm{v}}(x,y) & -i\delta_y \omega^{\mathrm{v}}(x,y) & 0 & 0 \\ -i\delta_x \omega^{\mathrm{v}}(x,y) & 0 & 0 & 0 \\ 0 & 0 & 0 & i\omega^{\mathrm{s}}(x,y) \\ 0 & 0 & -i\omega^{\mathrm{s}}(x,y) & 0 \end{pmatrix}. \tag{301}$$

Our definitions imply the commutation relations (239), (246) (with obvious modification to accommodate the Lie-algebra indices on the fields $A^I, C^I, \bar{C}^I$), as well as

$$A^I(x) \star_\hbar B^J(y) - B^J(y) \star_\hbar A^I(x) = \hbar k^{IJ} \delta_y \Delta^{\mathrm{v}}(x,y) \, 1\!\!1 \,. \tag{302}$$

The commutators of all other fields, in particular those involving any of the background fields $A_I^\ddagger, B_I^\ddagger, C_I^\ddagger, \bar{C}_I^\ddagger$, vanish. In this sense the background fields are $\mathbb{C}$-numbers, and their product is not deformed. This completes our construction of the quantization algebra $\mathbf{W}_0$ of free gauge theory.

The next step is to define within $\mathbf{W}_0$ the Wick products and time ordered products satisfying conditions T1–T11. As for the time ordered products with one factor, we make the same definition as in the scalar case, with the only difference that H is replaced by the matrix valued Hadamard parametrix

$$\big(H_{jk}(x,y)\big) = (k_{IJ}) \otimes \begin{pmatrix} H^{\mathrm{v}}(x,y) & -i\delta_y H^{\mathrm{v}}(x,y) & 0 & 0 \\ -i\delta_x H^{\mathrm{v}}(x,y) & 0 & 0 & 0 \\ 0 & 0 & 0 & iH^{\mathrm{s}}(x,y) \\ 0 & 0 & -iH^{\mathrm{s}}(x,y) & 0 \end{pmatrix}, \qquad (303)$$

where $j,k = (A^I, B^I, C^I, \bar{C}^I)$. Using the Hadamard parametrix, the time ordered products $T_1(\mathcal{O})$ with one factor $\mathcal{O} \in \mathbf{P}$ are defined by complete analogy with the scalar case, and they satisfy T1—T11. In particular, it follows from the definition that the Wick product $T_1(\mathbf{J}_0)$ of the free BRST-current (77) is conserved, $dT_1(\mathbf{J}_0) = T_1(d\mathbf{J}_0) = 0$ (modulo $\mathcal{J}_0$). Hence, we can define a a conserved BRST-charge (when the Cauchy surfaces are compact, see above). It also follows directly from the relations in the algebra $\mathbf{W}_0$ that $Q_0^2 = 0$ modulo $\mathbf{J}$. Thus, we can define the algebra of physical observables, $\hat{\mathcal{F}}_0$, by the cohomology of Q_0 as explained in the previous section. It follows from the Ward identity (c) below that if $\mathcal{O} \in \mathbf{P}$ is a classically gauge invariant polynomial expression in A^I, i.e., $\mathcal{O} = \prod \nabla^{s_i} dA^{I_i}$ (so that in particular $\hat{s}_0\mathcal{O} = 0$), then the corresponding Wick power $T_1(\mathcal{O})$ is in the kernel of Q_0 under the graded commutator. Thus, at ghost number 0, the algebra contains all local covariant quantum Wick powers of classically gauge invariant observables.

Thus, it only remains to prove the existence of Hadamard 2-point functions $\omega^{\mathrm{s}}, \omega^{\mathrm{v}}$ satisfying eq. (298), and to prove that the algebra $\hat{\mathcal{F}}_0$ has sensible Hilbert space representations. Both statements will now be proved by appealing to a deformation argument, as originally proposed by Fulling, Narcowich and Wald [49], and a construction of Fewster and Pfenning [46] for Maxwell fields on ultra-static spacetimes. That construction only works for spacetimes $M = \Sigma \times \mathbb{R}$ with Σ compact and simply connected (i.e., $H^1(\Sigma, d_\Sigma) = 0$), which is a physically reasonable assumption in view of the topological censorship theorem [52], and which we shall assume here.

Consider, besides the original spacetime, (M,g), an auxiliary deformed asymptotically static spacetime $(\hat{M},\hat{g})$. By this we mean that both spacetimes are identical to the future of some Cauchy surface $\Sigma \times \{t_+\}$, and that $\hat{g}$ is "ultrastatic" to the past of some Cauchy surface $\Sigma \times \{t_-\}$, meaning that $\hat{g}$ has the form

$$\hat{g} = -dt^2 + h(d\mathbf{x}, d\mathbf{x}) \qquad (304)$$

there, where $h = h_{ij}dx^i dx^j$ is a Riemannian metric on Σ that does not depend upon t. The idea of the deformation argument is now as follows. First, construct a pair $(\hat{\omega}^{\mathrm{s}}, \hat{\omega}^{\mathrm{v}})$ satisfying

the desired eq. (298) in the ultrastatic part of $(\hat{M},\hat{g})$. Then, because d and δ intertwine the action of the wave operators δd on 0-forms and $d\delta+\delta d$ on 1-forms, and since $(\hat{\omega}^{\rm s},\hat{\omega}^{\rm v})$ are bisolutions to the respective wave equations (297), the desired equations (298) must therefore hold on all of $(\hat{M},\hat{g})$, and not just on the ultrastatic part. Furthermore, one can show [81] using the celebrated "propagation of singularities theorem" [28] (see Appendix C,E) that the pair $(\hat{\omega}^{\rm s},\hat{\omega}^{\rm v})$ satisfies the desired wave front set condition (119) on all of $(\hat{M},\hat{g})$ if they are satisfied in the ultrastatic part. In particular, on the part of $(\hat{M},\hat{g})$ identical to (M,g), we then have a pair of Hadamard bi-distributions $(\hat{\omega}^{\rm s},\hat{\omega}^{\rm v})$ with the desired properties. The pair $(\hat{\omega}^{\rm s},\hat{\omega}^{\rm v})$ on the part of $(\hat{M},\hat{g})$ identical with (M,g) may now be propagated to a solution $(\omega^{\rm s},\omega^{\rm v})$ of the hyperbolic equations (297) on the undeformed spacetime (M,g). By the same arguments as above, this will now have a wave front set of Hadamard form on the undeformed spacetime, and it will satisfy the desired equation (298).

Thus, we need only prove the existence of a pair $(\hat{\omega}^{\rm s},\hat{\omega}^{\rm v})$ satisfying (298), the Hadamard condition (119), the commutator property, and field equations (297) on an ultrastatic spacetime $(\hat{M},\hat{g})$. This can be shown as follows using the following construction by [46], which in turn builds on results of [78]: On the 3-dimensional compact Riemannian spacetime (Σ,h), we consider a complete set of eigenfunctions of the corresponding scalar Laplace-operator $\Delta_h=d_\Sigma\delta_\Sigma$,

$$\Delta_h\varphi_{\mathbf{k}}=-\nu(S,\mathbf{k})^2\,\varphi_{\mathbf{k}}\,, \tag{305}$$

with positive eigenvalues $\nu(S,\mathbf{k})^2$, labelled by an index $\mathbf{k}\in J(S)$ in a corresponding index set. One defines $x=(t,\mathbf{x})\in\hat{M}$ and $u_{\mathbf{k}}(t,\mathbf{x})=e^{i\nu(S,\mathbf{k})t}\varphi_{\mathbf{k}}(\mathbf{x})$, as well as the "scalar" and "longitudinal" mode 1-forms on M by

$$\begin{aligned}\mathcal{A}_{S,\mathbf{k}}(t,\mathbf{x}) &= u_{\mathbf{k}}(t,\mathbf{x})\,dt & (306)\\ \mathcal{A}_{L,\mathbf{k}}(t,\mathbf{x}) &= \frac{1}{\nu(L,\mathbf{k})}\,du_{\mathbf{k}}(t,\mathbf{x})+iu_{\mathbf{k}}(t,\mathbf{x})\,dt\,, & (307)\end{aligned}$$

with $\nu(L,\mathbf{k})=\nu(S,\mathbf{k})$. These mode functions are smooth by elliptic regularity. One next chooses an orthonormal set of eigenmodes for the Laplacian $\Delta_h=d_\Sigma\delta_\Sigma+\delta_\Sigma d_\Sigma$ on (Σ,h) acting on 1-forms. By the Hodge decomposition theorem (see e.g. [13]), using $H^1(\Sigma,d_\Sigma)=0$, these can be uniquely decomposed into ones in the image of δ_Σ and those in the image of d_Σ. We denote those in the image of δ_Σ by $\xi_{\mathbf{k}}$ and their eigenvalues[13] by $-\nu(T,\mathbf{k})^2$, where $\mathbf{k}$ is now an index from a set $J(T)$. We define the corresponding "transversal" mode 1-forms on M by

$$\mathcal{A}_{T,\mathbf{k}}(t,\mathbf{x})=\mathrm{e}^{i\nu(T,\mathbf{k})t}\,\xi_{\mathbf{k}}(\mathbf{x})\,. \tag{308}$$

and we define the vector Hadamard 2-point distribution on the ultra-static spacetime by

$$\hat{\omega}^{\rm v}(x,y)=-\sum_{\lambda}\sum_{\mathbf{k}\in J(\lambda)}\frac{s(\lambda)}{2\nu(\lambda,\mathbf{k})}\mathcal{A}_{\lambda,\mathbf{k}}(x)\overline{\mathcal{A}_{\lambda,\mathbf{k}}(y)} \tag{309}$$

[13] Note that the scalar and transversal eigenvalues need not coincide.

where $s(S)=1, s(L)=-1=s(T)$, and $\lambda\in\{S,L,T\}$. It was proved in [46] that this is of Hadamard form and that it has the desired commutator property. We define the scalar Hadamard 2-point distribution on the ultra-static spacetime by

$$\hat{\omega}^{s}(x,y)=\sum_{\mathbf{k}\in J(S)}\frac{1}{2\nu(S,\mathbf{k})}u_{\mathbf{k}}(x)\overline{u_{\mathbf{k}}(y)}\,. \tag{310}$$

It was shown by [78] that this is of Hadamard form and that it satisfies the desired commutator property. The desired consistency property (298) on the ultrastatic spacetime follows by going through the definitions. Thus, by the deformation argument, we obtain from this a pair (ω^{v},ω^{s}) on the undeformed spacetime satisfying also the desired consistency condition (298).

We must finally construct a Hilbert space representation of the algebra $\mathcal{F}_0=\mathbf{W}_0/\mathcal{I}_0$ that gives rise to a corresponding representation of the algebra of physical observables (277) on the factor space (278). On an ultrastatic spacetime $\hat{M},\hat{g}$, we construct a representation as follows. We let $\mathfrak{h}_b$ be the 1-particle indefinite inner product space spanned by the orthonormal basis elements $e_{I,\lambda,\mathbf{k}}$, with $\lambda=S,L,T$ and $\mathbf{k}\in J(\lambda)$, with indefinite hermitian inner product defined by $(e_{I,\lambda,\mathbf{k}},e_{I',\lambda',\mathbf{k}'})=s(\lambda)k_{II'}\delta_{\lambda\lambda'}\delta_{\mathbf{k}\mathbf{k}'}$. We let

$$\mathfrak{F}_b=\bigoplus_{n=0}^{\infty}\bigotimes^{n}\mathfrak{h}_b \tag{311}$$

be the corresponding (indefinite metric) standard bosonic Fock space, with basis vectors

$$|I_1\lambda_1\mathbf{k}_1,\dots,I_n\lambda_n\mathbf{k}_n\rangle=\frac{1}{n!}\sum_{\pi\in S_n}e_{I_{\pi1}\lambda_{\pi1}\mathbf{k}_{\pi1}}\otimes\cdots\otimes e_{I_{\pi n}\lambda_{\pi n}\mathbf{k}_{\pi n}} \tag{312}$$

and we let $a^{+}_{I,\lambda,\mathbf{k}}$ be the standard creation operators associated with the basis vectors, i.e.,

$$a^{+}_{J,\nu,\mathbf{p}}|I_1\lambda_1\mathbf{k}_1,\dots,I_n\lambda_n\mathbf{k}_n\rangle=|J\nu\mathbf{p},I_1\lambda_1\mathbf{k}_1,\dots,I_n\lambda_n\mathbf{k}_n\rangle\,. \tag{313}$$

We let $\mathfrak{h}_f$ be the 1-particle indefinite inner product space spanned by the orthonormal basis elements $f_{I,\pm,\mathbf{k}}$ and $\mathbf{k}\in J(S)$, with indefinite hermitian inner product defined by $(f_{I,s,\mathbf{k}},f_{I',s',\mathbf{k}'})=i\varepsilon_{ss'}k_{II'}\delta_{\mathbf{k}\mathbf{k}'}$, where $\varepsilon_{ss'}$ is the anti-symmetric tensor in 2 dimensions. We let

$$\mathfrak{F}_f=\bigoplus_{n=0}^{\infty}\bigwedge^{n}\mathfrak{h}_f \tag{314}$$

be the corresponding (indefinite metric) standard fermionic Fock space, with basis vectors

$$|I_1s_1\mathbf{k}_1,\dots,I_ns_n\mathbf{k}_n\rangle=\frac{1}{n!}\sum_{\pi\in S_n}\mathrm{sgn}(\pi)\,f_{I_{\pi1}s_{\pi1}\mathbf{k}_{\pi1}}\otimes\cdots\otimes f_{I_{\pi n}s_{\pi n}\mathbf{k}_{\pi n}} \tag{315}$$

and we let $c^+_{I,s,\mathbf{k}}$ be the standard creation operators associated with the basis vectors, i.e.,

$$c^+_{J,r,\mathbf{p}}|I_1 s_1 \mathbf{k}_1, \dots, I_n s_n \mathbf{k}_n\rangle = |Jr\mathbf{p}, I_1 s_1 \mathbf{k}_1, \dots, I_n s_n \mathbf{k}_n\rangle\,. \tag{316}$$

The (indefinite) metric space $\mathcal{H}_0$ is defined as the tensor product $\mathcal{H}_0 = \mathfrak{F}_b \otimes \mathfrak{F}_f$. We now define the representatives of the fields $\Phi = (A^I, B^I, C^I, \bar{C}^I)$ as linear operators on $\mathcal{H}_0$ by

$$\pi_0(A^I(x)) = \sum_\lambda \sum_{\mathbf{k}\in J(\lambda)} \frac{1}{\sqrt{2\nu(\lambda,\mathbf{k})}} \mathcal{A}_{\lambda,\mathbf{k}}(x)\, a^+_{I,\lambda,\mathbf{k}} \quad + \text{h.c.} \tag{317}$$

$$\pi_0(C^I(x)) = \sum_{\mathbf{k}\in J(S)} \frac{1}{\sqrt{2\nu(S,\mathbf{k})}} u_\mathbf{k}(x) c^+_{I,+,\mathbf{k}} \quad + \text{h.c.} \tag{318}$$

$$\pi_0(\bar{C}^I(x)) = \sum_{\mathbf{k}\in J(S)} \frac{1}{\sqrt{2\nu(S,\mathbf{k})}} u_\mathbf{k}(x) c^+_{I,-,\mathbf{k}} \quad + \text{h.c.} \tag{319}$$

We define the representative $\pi_0(B^I(x))$ to be $-i\pi_0(\delta A^I(x))$, and we define the representative of any anti-field $\Phi^\ddagger$ to be zero. Finally, we define the representative of any element $F(u)$ of the form (299) by applying a normal ordering on the representatives (all creation operators to the left or all annihilation operators). The two-point functions of the vector- and ghost fields are then precisely given by $\hat{\omega}^{\mathrm{v}}$, resp. by $\hat{\omega}^{\mathrm{s}}$. As in flat spacetime, it may next be checked that, for compact G (i.e., positive definite Cartan-Killing form k_{IJ}) and in the ghost number 0 sector, the positivity requirement of sec. 4.2 is fulfilled. Thus, the physical Hilbert space (278) inherits a positive definite inner product. Furthermore, it follows from the consistency condition (298) that it contains precisely excitation of the longitudinal modes (308). In a general, non-static spacetimes, a similar construction can be applied by promoting the mode functions $\mathcal{A}_{\lambda,\mathbf{k}}, u_\mathbf{k}$ to solutions of the corresponding wave equation on the spacetime (M,g) by a deformation argument as above.

We expect a similar construction to work in the case when $H^1(\Sigma, d_\Sigma) \neq 0$, the only difference being the addition of corresponding zero modes to the mode expansions. We also expect a similar argument to work in spacetimes with non-compact Cauchy-surface, but it appears that this requires more work in general.

4.3 Interacting gauge theory

In this section, we describe in detail how the general construction of interacting Yang-Mills theory outlined in sec. 4.1 is performed. To construct perturbatively the interacting fields in interacting gauge theory, we need to construct the time-ordered products in the free theory considered in the previous subsection. For time ordered products with 1 factor, this was done there. For time ordered products with n factors, this can be done as described in Sect. 3, and these time ordered products will satisfy the analog of conditions T1–T11.

However, in gauge theory, the time ordered products must satisfy further constraints related to gauge invariance. As we have argued in section 4.2, in the gauge fixed formalism, we need

to be able to define an interacting BRST-charge operator, Q_I, and we need that operator to be nilpotent, i.e. $Q_I^2 = 0$. In order to meaningfully construct Q_I, we need a conserved interacting BRST-current $\mathbf{J}_I$. If our time ordered products only satisfy T1–T11 [with the symmetry property T6 replaced by graded symmetry with respect to the Grassmann parity], then there is in general no guarantee that the interacting BRST-current is conserved, $d\mathbf{J}_I = 0$, nor that $Q_I^2 = 0$, nor that $[Q_I, \Psi_I] = 0$ for strictly gauge invariant operators Ψ of ghost number 0.

We will now formulate a set of Ward identities in the free theory that will guarantee that these conditions are satisfied, and which moreover will guarantee (formally) that the S-matrix—when it exists—is BRST-invariant. As argued in the previous section, with such a definition of time-ordered products, the conditions of gauge invariance of the perturbative interacting quantum field theory are then satisfied. The Ward identities that we want to propose are to be viewed as an additional normalization condition on the time ordered product, and are as follows. Consider a local operator $\mathcal{O} \in \mathbf{P}$, given by an expansion of the form

$$\mathcal{O} = \mathcal{O}_0 + \lambda \mathcal{O}_1 + \dots \lambda^N \mathcal{O}_N \,. \tag{320}$$

Let f be a smooth compactly supported test function on M, and let

$$F = \int_M [\mathcal{O}_0 + \lambda f \mathcal{O}_1 + \dots + \lambda^N f^N \mathcal{O}_N] \,. \tag{321}$$

Then the Ward identity that we will consider is

$$\boxed{\left[Q_0, T\left(\mathrm{e}_\otimes^{iF/\hbar}\right)\right] = -\frac{1}{2} T\left((S_0 + F, S_0 + F) \otimes \mathrm{e}_\otimes^{iF/\hbar}\right) \quad \text{modulo } \mathcal{J}_0 \,.} \tag{322}$$

Here, Q_0 is the free BRST-charge operator, $(\,.\,,\,.\,)$ is the anti-bracket (58), and $[\,,\,]$ is the graded commutator in the algebra $\mathbf{W}_0$. As with all generating type formulae in this work, this is to be understood as a shorthand for the hierarchy of identities that are obtained when the above expression is expanded as a formal power series in λ. We now write out explicitly this hierarchy of identities. For this, it is convenient to introduce some notation. We denote by $I = \{k_1, \dots, k_r\}$ subsets of $\underline{n} = \{1, \dots, n\}$, and we write $r = |I|$ for the number of elements. We set $X_I = (x_{k_1}, \dots, x_{k_r})$, and we put

$$\mathcal{O}_r(X_I) = r!\, \mathcal{O}_r(x_{k_1}) \delta(x_{k_1}, \dots, x_{k_r}). \tag{323}$$

With these notations, the Ward-identity (322) can be expressed as

$$\begin{aligned}
&\sum_{I_1 \cup \dots \cup I_t = \underline{n}} \left(\frac{i}{\hbar}\right)^t \left[Q_0, T_t(\mathcal{O}_{|I_1|}(X_{I_1}) \otimes \dots \mathcal{O}_{|I_t|}(X_{I_t}))\right] = \\
&\qquad - \sum_{I_1 \cup \dots \cup I_t = \underline{n}} \left(\frac{i}{\hbar}\right)^{t-1} \sum_{k=1}^{t} (-1)^{\varepsilon_k} T_t(\mathcal{O}_{|I_1|}(X_{I_1}) \otimes \dots \hat{s}_0 \mathcal{O}_{|I_k|}(X_{I_k}) \otimes \dots \mathcal{O}_{|I_t|}(X_{I_t})) \\
&- \sum_{I_1 \cup \dots \cup I_t = \underline{n}} \left(\frac{i}{\hbar}\right)^{t-2} \sum_{1 \le k < l \le n} (-1)^{\varepsilon_k \varepsilon_l} T_{t-1}(\mathcal{O}_{|I_1|}(X_{I_1}) \otimes \dots (\mathcal{O}_{|I_k|}(X_{I_k}), \mathcal{O}_{|I_l|}(X_{I_l})) \otimes \dots \mathcal{O}_{|I_t|}(X_{I_t}))
\end{aligned} \tag{324}$$

modulo $\mathcal{J}_0$, where $\varepsilon_k = \varepsilon(\mathcal{O}_1) + \cdots + \varepsilon(\mathcal{O}_{k-1})$. We will not prove the above Ward identities for arbitrary operators $\mathcal{O}$ in this work, but only for certain special cases, which are relevant for our analysis of gauge invariance. These cases are

(T12a) $\mathcal{O}$ is given by the interaction Lagrangian, $\mathcal{O} = \lambda \mathbf{L}_1 + \lambda^2 \mathbf{L}_2$,

(T12b) $\mathcal{O}$ is given by a linear combination of the interaction Lagrangian, and the BRST-current $\mathcal{O} = \lambda \mathbf{L}_1 + \lambda^2 \mathbf{L}_2 + \gamma \wedge (\mathbf{J}_0 + \lambda \mathbf{J}_1)$ (evaluation of the Ward identity to first order in $\gamma \in \Omega^1_0(M)$).

(T12c) $\mathcal{O} = \lambda \mathbf{L}_1 + \lambda^2 \mathbf{L}_2 + \gamma \wedge \sum \lambda^k \Psi_k \in \mathbf{P}^4(M)$ is given by a linear combination of the interaction Lagrangian and a strictly gauge invariant operator $\Psi = \sum_k \lambda^k \Psi_k \in \mathbf{P}^p(M)$ of ghost number 0, i.e., of the form given by eq. (47) (evaluation of the Ward identity to first order in $\gamma \in \Omega^{4-p}_0(M)$).

It is only for those cases that we will prove the Ward-identities (324), and that proof is provided in section 4.4. For convenience, we now give explicitly the form of the Ward-identities in the cases (a), (b), and (c).

Case (T12a) The Ward identities in that case are given explicitly by

$$
\begin{aligned}
\sum_{I_1 \cup \cdots \cup I_t = \underline{n}} \left(\frac{i}{\hbar}\right)^t \left[Q_0, T_t(\mathbf{L}_{|I_1|}(X_{I_1}) \otimes \dots \mathbf{L}_{|I_t|}(X_{I_t}))\right] = \\
- \sum_{I_1 \cup \cdots \cup I_t = \underline{n}} \left(\frac{i}{\hbar}\right)^{t-1} \sum_{k=1}^{t} T_t(\mathbf{L}_{|I_1|}(X_{I_1}) \otimes \dots \hat{s}_0 \mathbf{L}_{|I_k|}(X_{I_k}) \otimes \dots \mathbf{L}_{|I_t|}(X_{I_t})) \\
- \sum_{I_1 \cup \cdots \cup I_t = \underline{n}} \left(\frac{i}{\hbar}\right)^{t-2} \sum_{1 \le j < k \le t} T_{t-1}(\mathbf{L}_{|I_1|}(X_{I_1}) \otimes \dots (\mathbf{L}_{|I_j|}(X_{I_j}), \mathbf{L}_{|I_k|}(X_{I_k})) \otimes \dots \mathbf{L}_{|I_t|}(X_{I_t})),
\end{aligned}
\tag{325}
$$

modulo $\mathcal{J}_0$.

Case (T12b) The Ward identities in that case are given explicitly by

$$
\begin{aligned}
\sum_{I_1 \cup \cdots \cup I_t = \underline{n}} \left(\frac{i}{\hbar}\right)^{t-1} \left[Q_0, T_t(\mathbf{J}_{|I_1|}(y, X_{I_1}) \otimes \mathbf{L}_{|I_2|}(X_{I_2}) \otimes \cdots \otimes \mathbf{L}_{|I_t|}(X_{I_t}))\right] = \\
\sum_{I_1 \cup \cdots \cup I_t = \underline{n}} \left(\frac{i}{\hbar}\right)^{t-2} \sum_{i=2}^{t} T_t(\mathbf{J}_{|I_1|}(y, X_{I_1}) \otimes \mathbf{L}_{|I_2|}(X_{I_2}) \otimes \dots \hat{s}_0 \mathbf{L}_{|I_i|}(X_{I_i}) \otimes \dots \mathbf{L}_{|I_t|}(X_{I_t})) \\
- \sum_{I_1 \cup \cdots \cup I_t = \underline{n}} \left(\frac{i}{\hbar}\right)^{t-2} T_t(\hat{s}_0 \mathbf{J}_{|I_1|}(y, X_{I_1}) \otimes \mathbf{L}_{|I_2|}(X_{I_2}) \otimes \cdots \otimes \mathbf{L}_{|I_t|}(X_{I_t})) \\
+ \sum_{I_1 \cup \cdots \cup I_t = \underline{n}} \left(\frac{i}{\hbar}\right)^{t-3} \sum_{2 \le i < j \le t} T_{t-1}(\mathbf{J}_{|I_1|}(y, X_{I_1}) \otimes \mathbf{L}_{|I_2|}(X_{I_2}) \otimes \dots (\mathbf{L}_{|I_i|}(X_{I_i}), \mathbf{L}_{|I_j|}(X_{I_j})) \otimes \dots \mathbf{L}_{|I_t|}(X_{I_t})) \\
- \sum_{I_1 \cup \cdots \cup I_t = \underline{n}} \left(\frac{i}{\hbar}\right)^{t-2} \sum_{2 \le i \le t} T_{t-1}(\mathbf{L}_{|I_2|}(X_{I_2}) \otimes \dots (\mathbf{J}_{|I_1|}(y, X_{I_1}), \mathbf{L}_{|I_i|}(X_{I_i})) \otimes \dots \mathbf{L}_{|I_t|}(X_{I_t})),
\end{aligned}
\tag{326}
$$

modulo $\mathcal{J}_0$. Here $\mathbf{J}_1(y,x) = \mathbf{J}_1(y)\delta(x,y)$.

Case (T12c) Let $\Psi = \Psi_0 + \lambda\Psi_1 + \cdots + \lambda^N \Psi_N$ be a strictly gauge invariant local field polynomial of ghost number zero. Thus, by formula (47), up to local curvature terms which we may ignore, $\Psi = \prod \Theta_{s_i}(F, \mathcal{D}F, \mathcal{D}^2 F, \dots)$, where Θ_s are invariant polynomials of the Lie-algebra. The Ward identities in that case are given explicitly by

$$
\begin{aligned}
&\sum_{I_1\cup\cdots\cup I_t=\underline{n}} \left(\frac{i}{\hbar}\right)^{t-1} \left[Q_0, T_t(\Psi_{|I_1|}(y,X_{I_1}) \otimes \mathbf{L}_{|I_2|}(X_{I_2}) \otimes \cdots \otimes \mathbf{L}_{|I_t|}(X_{I_t}))\right] = \\
&\quad - \sum_{I_1\cup\cdots\cup I_t=\underline{n}} \left(\frac{i}{\hbar}\right)^{t-2} \sum_{i=2}^{t} T_t(\Psi_{|I_1|}(y,X_{I_1}) \otimes \mathbf{L}_{|I_2|}(X_{I_2}) \otimes \dots \hat{s}_0 \mathbf{L}_{|I_i|}(X_{I_i}) \otimes \dots \mathbf{L}_{|I_t|}(X_{I_t})) \\
&\qquad - \sum_{I_1\cup\cdots\cup I_t=\underline{n}} \left(\frac{i}{\hbar}\right)^{t-2} T_t(\hat{s}_0 \Psi_{|I_1|}(y,X_{I_1}) \otimes \mathbf{L}_{|I_2|}(X_{I_2}) \otimes \cdots \otimes \mathbf{L}_{|I_t|}(X_{I_t})) \\
&- \sum_{I_1\cup\cdots\cup I_t=\underline{n}} \left(\frac{i}{\hbar}\right)^{t-3} \sum_{2\le i<j\le t} T_{t-1}(\Psi_{|I_1|}(y,X_{I_1}) \otimes \mathbf{L}_{|I_2|}(X_{I_2}) \otimes \dots (\mathbf{L}_{|I_i|}(X_{I_i}), \mathbf{L}_{|I_j|}(X_{I_j})) \otimes \dots \mathbf{L}_{|I_t|}(X_{I_t})) \\
&\quad - \sum_{I_1\cup\cdots\cup I_t=\underline{n}} \left(\frac{i}{\hbar}\right)^{t-2} \sum_{i=2}^{t} T_{t-1}(\mathbf{L}_{|I_2|}(X_{I_2}) \otimes \dots (\Psi_{|I_1|}(y,X_{I_1}), \mathbf{L}_{|I_i|}(X_{I_i})) \otimes \dots \mathbf{L}_{|I_t|}(X_{I_t}))\,,
\end{aligned} \tag{327}
$$

modulo $\mathcal{J}_0$.

We will give a proof of the Ward-identities T12a–T12c in subsec. 4.4. We will then show in subsec. 4.6 that the Ward identities T12a imply the conservation of the interacting BRST-current, $d\mathbf{J}_I = 0$. We will prove in subsec. 4.7 that the Ward identities T12b furthermore imply that $Q_I^2 = 0$ and we will show in subsec. 4.8 that the Ward identities T12c imply $[Q_I, \Psi_I] = 0$ for strictly gauge invariant operators Ψ at ghost number 0. The Ward identity T12a also formally implies the BRST-invariance of the S-matrix (see subsec. 4.5), provided the latter exists (which is not the case in Minkowski space, and appears even more unlikely in curved spacetime). We will not analyze this existence question here, so in this sense the BRST-invariance of the S-matrix is not a rigorous result unlike the other results in our paper.

As an aside, we note that, the Ward identities T12a, T12b, and T12c are *incompatible* with the identity

$$
\begin{aligned}
&\left[Q_0, T_n(\mathcal{O}_1(x_1) \otimes \cdots \otimes \mathcal{O}_n(x_n))\right] \\
&\qquad = i\hbar \sum_{i=1}^{n} (-1)^{\varepsilon_i} T_n(\mathcal{O}_1(x_1) \otimes \cdots \hat{s}_0 \mathcal{O}_i(x_i) \otimes \cdots \mathcal{O}_n(x_n)) \quad \text{mod } \mathcal{J}_0 \text{ (WRONG!)}\,,
\end{aligned} \tag{328}
$$

unless none of the fields $\mathcal{O}_i$ contains anti-fields. The above identity has been considered before in the context of flat spacetime in [36], where it has been termed “Master BRST-identity.” It appears that it is impossible to satisfy this identity (even for $n = 1$) when anti-fields are present. It would also not imply either the conservation of the interacting BRST current $\mathbf{J}_I$ nor

the nilpotency of the interacting BRST charge in a framework with anti-fields. Since the use of anti-fields also appears to be essential in order to derive sufficiently strong constraints on potential anomalies to the BRST-Ward identities, we believe that eq. (328) is not a good starting point for the proof of gauge invariance in perturbative Yang-Mills theory.

4.4 Inductive proof of Ward identities T12a, T12b, and T12c

We now show that the Ward identities can be satisfied together with T1—T11 by making a suitable redefinition of the time-ordered products if necessary. The Ward identity (324) is an identity modulo $\mathcal{J}_0$, that is, it is required to hold only on shell. For the proof of that identity it is actually useful to consider a more stringent "off-shell" version of the identity. Even though that off-shell version is more stringent, it will in fact turn out to be easier to prove, as it gives, at the same time, stronger constraints of cohomological nature on the the possible anomalies than the corresponding on-shell version.

To set up the off-shell version of our Ward-identity, we first recall the definition $\hat{s}_0 = s_0 + \sigma_0$ of the free Slavnov-Taylor differential, given above in eq. (71) and (72). As it stands, the differential $\hat{s}_0$ was defined as a map $\hat{s}_0 : \mathbf{P}(M) \to \mathbf{P}(M)$, i.e., it acts on polynomial expressions in the *classical* fields $\Phi, \Phi^\ddagger$. We will now extend the action of $\hat{s}_0$ to the non-commutative algebra $\mathbf{W}_0$. For this, we recall that the algebra $\mathbf{W}_0$ may be viewed as the closure of the CCR-algebra $\mathbf{W}_{00}$, which in turn is generated by expressions of the form $F_1 \star_\hbar \dots \star_\hbar F_n$, where each F_i is given by $\int f_i \wedge \mathcal{O}_i$, with f_i smooth and of compact support, and with $\mathcal{O}_i$ given by one of the "basic fields" $\Phi, \Phi^\ddagger$. To define the action of $\hat{s}_0$ on such elements of $\mathbf{W}_{00}$, we set

$$\hat{s}_0\Big(\mathcal{O}_1(x_1) \star_\hbar \dots \star_\hbar \mathcal{O}_s(x_n)\Big) = \sum_{i=1}^{n} (-1)^{\varepsilon_i} \mathcal{O}_1(x_1) \star_\hbar \dots \hat{s}_0 \mathcal{O}_i(x_i) \star_\hbar \dots \mathcal{O}_n(x_n)\,, \tag{329}$$

where $\mathcal{O}_i$ is either a basic field Φ, or an anti-field $\Phi^\ddagger$. This defines Slavnov-Taylor differential $\hat{s}_0$ as a graded derivation (denoted by the same symbol) of the algebra $\mathbf{W}_{00}$. As we have remarked, the subalgebra $\mathbf{W}_{00} \subset \mathbf{W}_0$ is dense (in the Hörmander topology). Thus, we can uniquely extend $\hat{s}_0$ to a graded derivation on $\mathbf{W}_0$ by continuity with respect to this topology. We will again denote this graded derivation $\hat{s}_0 : \mathbf{W}_0 \to \mathbf{W}_0$ by the same symbol. Actually, we must still check that the definition (329) is consistent, i.e., compatible with the algebra relations in $\mathbf{W}_{00}$. We formulate this result as a lemma:

Lemma 7. The formula (329) defines a graded derivation on $\mathbf{W}_0$.

Proof: The basic algebraic relations in $\mathbf{W}_{00}$ are the graded commutation relations

$$[\Phi^i(x), \Phi^j(y)] = i\hbar\Delta^{ij}(x,y)\,\mathbb{1}\,, \quad [\Phi^i(x), \Phi^\ddagger_j(y)] = 0 = [\Phi^\ddagger_i(x), \Phi^\ddagger_j(y)]\,, \tag{330}$$

where Δ^{ij} is the matrix of commutator functions given by

$$\left(\Delta^{jk}(x,y)\right) = (k_{IJ}) \otimes \begin{pmatrix} \Delta^{\mathrm{v}}(x,y) & -i\delta_y\Delta^{\mathrm{v}}(x,y) & 0 & 0 \\ -i\delta_x\Delta^{\mathrm{v}}(x,y) & 0 & 0 & 0 \\ 0 & 0 & 0 & i\Delta^{\mathrm{s}}(x,y) \\ 0 & 0 & -i\Delta^{\mathrm{s}}(x,y) & 0 \end{pmatrix}, \tag{331}$$

where $\Phi^i = (A^I, B^I, C^I, \bar{C}^I)$, and where $\Delta^{\mathrm{v}}, \Delta^{\mathrm{s}}$ are the advanced minus retarded propagators for vectors and scalars, see Appendix E. To show that the definition of $\hat{s}_0$ on $\mathbf{W}_{00}$ is consistent, we next apply the definition (329) to the above graded commutators and check that we get identities. This follows from the relations

$$d_x\Delta^{\mathrm{s}}(x,y) = -\delta_y\Delta^{\mathrm{v}}(x,y), \quad d_y\Delta^{\mathrm{s}}(x,y) = -\delta_x\Delta^{\mathrm{v}}(x,y), \tag{332}$$

which in turn a direct consequence of the field equations satisfied by the advanced and retarded propagators for scalars and vectors. $\square$

We are now in a position to formulate the desired off-shell version of our (anomalous) Ward identity that will eventually enable us to prove T12a, T12b, and T12c. We formulate our result in a proposition:

Proposition 3: (Anomalous Ward Identity) For a general prescription for time-ordered products satisfying T1—T11, the identity

$$\boxed{\hat{s}_0 T\left(\mathrm{e}_\otimes^{iF/\hbar}\right) = \frac{i}{2\hbar} T\Big((S_0 + F, S_0 + F) \otimes \mathrm{e}_\otimes^{iF/\hbar}\Big) + \frac{i}{\hbar} T\Big(A(\mathrm{e}_\otimes^F) \otimes \mathrm{e}_\otimes^{iF/\hbar}\Big)} \tag{333}$$

holds. Here $F = \int f \wedge \mathcal{O}$ is any smeared local field with $\mathcal{O} \in \mathbf{P}^p(M)$, $f \in \Omega_0^{4-p}(M)$ and $A(\mathrm{e}_\otimes^F)$ is the anomaly, given by

$$A(\mathrm{e}_\otimes^F) = \sum_{n \geq 0} \frac{1}{n!} A_n(F^{\otimes n}), \tag{334}$$

and where $A_n : \mathbf{P}^{k_1}(M) \otimes \cdots \otimes \mathbf{P}^{k_n}(M) \to \mathbf{P}^{k_1/\dots/k_n}(M^n)$ are local functionals supported on the total diagonal. The anomaly satisfies the following further properties:

(i) $A(\mathrm{e}_\otimes^F) = O(\hbar)$.

(ii) Each A_n is locally and covariantly constructed out of the metric.

(iii) Each A_n has ghost number one, in the sense that $\mathcal{N}_g \circ A_n - A_n \circ \Gamma_n \mathcal{N}_g = A_n$, where $\mathcal{N}_g$ is the number counter for the ghost fields, see eq. (50) (with additional terms for the anti-fields), and

$$\Gamma_n \mathcal{N}_g = \sum_{i=1}^n id \otimes \cdots \mathcal{N}_g \otimes \cdots id : \mathbf{P}^{\otimes n} \to \mathbf{P}^{\otimes n}. \tag{335}$$

(iv) Each A_n has dimension 0, in the sense that $\mathcal{N}_d \circ A_n - A_n \circ \Gamma_n \mathcal{N}_d = 0$, where $\mathcal{N}_d := \mathcal{N}_f + \mathcal{N}_r$ is the dimension counter, which is the sum of the dimensions $\mathcal{N}_f$ of the individual fields and anti-fields (see the tables above), and the dimensions $\mathcal{N}_r$ of the curvature terms.

(v) The maps A_n are real in the sense that $A(\mathrm{e}_\otimes^F)^* = A(\mathrm{e}_\otimes^{F^*})$.

Before we come to the proof of this key proposition, we note that, in the absence of anomalies $A(\mathrm{e}_\otimes^F) = 0$, the off-shell version of our Ward-identity becomes

$$\boxed{\hat{s}_0 T\left(\mathrm{e}_\otimes^{iF/\hbar}\right) = \frac{i}{2\hbar} T\left((S_0+F, S_0+F) \otimes \mathrm{e}_\otimes^{iF/\hbar}\right).} \tag{336}$$

The difference to (322) is that on the left side, we do not have the graded commutator with Q_0, but instead we act with the Slavnov-Taylor map $\hat{s}_0$, which is the sum of the standard free BRST-differential s_0 generated by Q_0, and the Koszul-Tate differential. The addition of the Koszul-Tate differential is crucial to obtain an identity that holds *off shell*, and not just modulo the free field equations as eq. (322). As already indicated, despite being more stringent, the sharpened off-shell Ward identity (336) is in fact simpler to prove than the corresponding on-shell identity (322), as it also allows one to derive more stringent consistency conditions on the possible anomalies. These consistency conditions rely in an essential way upon the use of the anti-fields, and this is the principal reason why we have introduced such fields in our construction.

Proof of Proposition 3: The proof of the anomalous Ward-identity (333) proceeds by induction in the order n in perturbation theory, noting that the anomalous Ward-identity holds at order n if it holds up to order $n-1$, modulo a contribution supported on the total diagonal. That contribution is defined to be A_n. In more detail, consider n local functionals $F_1, \dots, F_n$ with $F_i = \int f_i \wedge \mathcal{O}_i$, with f_i a form of compact support and form degree complementary to that of $\mathcal{O}_i \in \mathbf{P}(M)$. For definiteness and simplicity, we assume that all F_i have Grassmann parity 0; in the general case one proceeds similarly. The anomalous Ward-identity (333) at order n is then the statement that

$$\begin{aligned}
&\hat{s}_0 T_n(F_1 \otimes \cdots \otimes F_n) = \\
&\sum_{k=0}^{n} T_n(F_1 \otimes \dots \hat{s}_0 F_k \otimes \dots F_n) + \frac{\hbar}{i} \sum_{k<j} T_{n-1}(F_1 \otimes \dots (F_j, F_k) \otimes \dots F_n) + \\
&\sum_{t=1}^{n} \sum_{k_1 < \cdots < k_t} \sum_{l_1 < \cdots < l_{n-t}} \left(\frac{\hbar}{i}\right)^{t-1} T_{n-t+1}(A_t(F_{k_1} \otimes \dots F_{k_t}) \otimes F_{l_1} \otimes \dots F_{l_{n-t}}).
\end{aligned} \tag{337}$$

We now look at the individual terms in this expression. We decompose $\hat{s}_0 = s_0 + \sigma_0$ into its pure BRST-part s_0 and the Koszul-Tate differential σ_0. Letting ε_i be the Grassmann parity of f_i

(equal to that of O_i, since F_i is assumed to be bosonic), we have

$$
\begin{aligned}
& \sigma_0 T_n(F_1 \otimes \cdots \otimes F_n) \\
= \;& \sigma_0 \Big((-1)^{\sum_{i<j} \varepsilon_i \varepsilon_j} \int f_1(x_1) \dots f_n(x_n) \, T_n(O_1(x_1) \otimes \cdots \otimes O_n(x_n)) \, dx_1 \dots dx_n \Big) \\
= \;& (-1)^{\sum_{i<j} \varepsilon_i \varepsilon_j} \int \sum_{k=1}^{n} (-1)^{\sum_{l<k} \varepsilon_l} [f_1(x_1) \dots \sigma_0 f_k(x_k) \dots f_n(x_n)] \star_\hbar T_n \big(\otimes_i O_i(x_i) \big) \, dx_1 \dots dx_n \\
= \;& (-1)^{\sum_{i<j} \varepsilon_i \varepsilon_j} \int \sum_{k=1}^{n} (-1)^{\sum_{l<k} \varepsilon_l} [f_1(x_1) \dots \frac{\delta_R S_0}{\delta \Phi(y)} \frac{\delta_L f_k(x_k)}{\delta \Phi^\dagger(y)} \dots f_n(x_n)] \star_\hbar T_n \big(\otimes_i O_i(x_i) \big) \, dy dx_1 \dots dx_n \\
= \;& \sum_{k=1}^{n} \int \frac{\delta_R S_0}{\delta \Phi(y)} \star_\hbar T_n \left(F_1 \otimes \dots \frac{\delta_L F_k}{\delta \Phi^\ddagger(y)} \otimes \dots F_n \right) dy \,, \qquad (338)
\end{aligned}
$$

and we have

$$
\sum_{k=1}^{n} T_n(F_1 \otimes \dots \sigma_0 F_k \otimes \dots F_n) = \sum_{k=1}^{n} \int T_n \left(F_1 \otimes \dots \frac{\delta_R S_0}{\delta \Phi(x)} \wedge \frac{\delta_L F_k}{\delta \Phi^\ddagger(x)} \otimes \dots F_n \right) dx \,, \qquad (339)
$$

using the definition of σ_0, see eq. (72) and the following table. We may combine these two identities into the following identity for the corresponding generating functionals:

$$
\begin{aligned}
& \sigma_0 T(\mathrm{e}_\otimes^{iF/\hbar}) - \frac{i}{\hbar} T(\sigma_0 F \otimes \mathrm{e}_\otimes^{iF/\hbar}) \qquad (340) \\
= \;& \frac{i}{\hbar} \int \frac{\delta_R S_0}{\delta \Phi(x)} \star_\hbar T \Big(\frac{\delta_L F}{\delta \Phi^\ddagger(x)} \otimes \mathrm{e}_\otimes^{iF/\hbar} \Big) dx - \frac{i}{\hbar} \int T \Big(\big(\frac{\delta_R S_0}{\delta \Phi(x)} \wedge \frac{\delta_L F}{\delta \Phi^\ddagger(x)} \big) \otimes \mathrm{e}_\otimes^{iF/\hbar} \Big) dx \,.
\end{aligned}
$$

To manipulate this expression, we now use a proposition formulated and proven first in [15] [see eq. (5.48) in lemma 11 of this reference].

Proposition 4: ("Master Ward Identity") Let $\psi \in C_0^\infty(M) \cdot \mathbf{P}(M)$ be arbitrary, i.e., ψ is a local functional of the fields, times a compactly supported cutoff function. Set

$$
B = \int_M \frac{\delta S_0}{\delta \Phi(x)} \wedge \psi(x) \,, \quad \delta_B F = \int_M \frac{\delta F}{\delta \Phi(x)} \wedge \psi(x) \,. \qquad (341)
$$

Then we have

$$
T \big(\big[F + \delta_B F + \Delta_B(\mathrm{e}_\otimes^F) \big] \otimes \mathrm{e}_\otimes^{iF/\hbar} \big) = \int \frac{\delta S_0}{\delta \Phi(x)} \star_\hbar T \big(\psi(x) \otimes \mathrm{e}_\otimes^{iF/\hbar} \big) \, dx \,. \qquad (342)
$$

Here $\Delta_B(\mathrm{e}_\otimes^F) = \sum_n \frac{1}{n!} \Delta_n(F^{\otimes n})$ and each $\Delta_n : \mathbf{P}^{k_1}(M) \otimes \cdots \otimes \mathbf{P}^{k_n}(M) \to \mathbf{P}^{k_1/\dots/k_n}(M^n)$ is a linear map that is supported on the total diagonal. If the F_i do not depend on $\hbar$, then the quantity $\Delta_n(F_1 \otimes \cdots \otimes F_n)$ is of order $O(\hbar)$.

We will outline the proof of this proposition at the end of the present proof. We now apply the Master Ward identity to the case when $\psi(x) = \delta_L F / \delta\Phi^\ddagger(x)$. Then we obtain, for the last term in eq. (340) the expression

$$
\begin{aligned}
&-\frac{i}{\hbar}\int T\Big(\big(\frac{\delta_R S_0}{\delta\Phi(x)} \wedge \frac{\delta_L F}{\delta\Phi^\ddagger(x)}\big) \otimes e_\otimes^{iF/\hbar}\Big) = \\
&\frac{i}{\hbar} T\left(\delta_B F \otimes e_\otimes^{iF/\hbar}\right) - \frac{i}{\hbar}\int \frac{\delta_R S_0}{\delta\Phi(x)} \star_\hbar T\left(\frac{\delta_L F}{\delta\Phi^\ddagger(x)} \otimes e_\otimes^{iF/\hbar}\right) + \\
&\frac{i}{\hbar} T\left(\Delta_B(e_\otimes^F) \otimes e_\otimes^{iF/\hbar}\right).
\end{aligned} \tag{343}
$$

Now, we have, with our choice $\psi(x) = \delta_L F / \delta\Phi^\ddagger(x)$,

$$
\delta_B F = \int_M \frac{\delta_R F}{\delta\Phi(x)} \wedge \frac{\delta_L F}{\delta\Phi^\ddagger(x)} = \frac{1}{2}(F,F). \tag{344}
$$

Thus, we altogether obtain the identity

$$
\boxed{\sigma_0 T\left(e_\otimes^{iF/\hbar}\right) - \frac{i}{\hbar} T\left(\sigma_0 F \otimes e_\otimes^{iF/\hbar}\right) = \frac{i}{2\hbar} T\left((F,F) \otimes e_\otimes^{iF/\hbar}\right) + \frac{i}{\hbar} T\left(\Delta_B(e_\otimes^F) \otimes e_\otimes^{iF/\hbar}\right)} \tag{345}
$$

which is in fact just another equivalent way of expressing the Master Ward Identity. This identity in effect will take care of all terms in eq. (337) involving the Koszul-Tate differential. We now look at the terms involving the pure BRST-differential s_0. To deal with these terms, we now use the following identity:

Lemma 8.

$$
\begin{aligned}
&s_0 T_n(F_1 \otimes \cdots \otimes F_n) = \sum_{k=0}^{n} T_n(F_1 \otimes \ldots s_0 F_k \otimes \ldots F_n) + \\
&\sum_{t=1}^{n} \sum_{k_1 < \cdots < k_t} \sum_{l_1 < \cdots < l_{n-t}} \left(\frac{\hbar}{i}\right)^{t-1} T_{n-t+1}(\delta_t(F_{k_1} \otimes \ldots F_{k_t}) \otimes F_{l_1} \otimes \ldots F_{l_{n-t}}).
\end{aligned} \tag{346}
$$

Here, δ_n is a map of the same nature as A_n, i.e., it is supported on the total diagonal, and it is of order $O(\hbar)$. A formula generating these identities is

$$
s_0 T\left(e_\otimes^{iF/\hbar}\right) - \frac{i}{\hbar} T\left(s_0 F \otimes e_\otimes^{iF/\hbar}\right) = \frac{i}{\hbar} T\left(\delta(e_\otimes^F) \otimes e_\otimes^{iF/\hbar}\right). \tag{347}
$$

Proof of Lemma 8: For $n = 1$ the identity says that $s_0 T_1(F) = T_1(s_0 F) + T_1(\delta_1(F))$, and we simply define $\delta_1(F)$ in this way. Since there is no anomaly in the classical limit, it follows that $\delta_1(F)$ is of order $\hbar$. We now proceed inductively to prove the equation for all n. Assume that it has been shown for any number of factors up to $n-1$, and the $\delta_1, \ldots, \delta_{n-1}$ have consequently

been defined. Take n functionals $F_1, \dots, F_n$ with the property that the support of the first l functionals is not in the future of the support of the last $n-l$ functionals, where l is not equal to 0 or n. Define M_n to be the difference between the left and right terms in the above equation, with the n-th term in the sum (the one containing δ_n) omitted. Then, using the causal factorization property of the time ordered products and the assumed support properties of the F_i, it follows that

$$\begin{aligned}
&M_n(F_1 \otimes \cdots \otimes F_n) = -s_0\big(T_l(F_1 \otimes \cdots \otimes F_l) \star_\hbar T_{n-l}(F_{l+1} \otimes \cdots \otimes F_n)\big) + \\
&\sum_{k=0}^{l} T_l(F_1 \otimes \dots s_0 F_k \otimes \dots F_l) \star_\hbar T_{n-l}(F_{l+1} \otimes \cdots \otimes F_n) + \\
&\sum_{k=l+1}^{l} T_n(F_1 \otimes \cdots \otimes \dots F_l) \star_\hbar T_{n-l}(F_{l+1} \otimes \dots s_0 F_k \otimes \dots F_n) + \\
&\sum_{t=1}^{l} \sum_{k_1 < \cdots < k_t \le l} \sum_{l_1 < \cdots < l_{n-t} \le l} \left(\frac{\hbar}{i}\right)^{t-1} T_{l-t+1}(\delta_t(F_{k_1} \otimes \dots F_{k_t}) \otimes F_{l_1} \otimes \dots F_{l_{l-t}}) \star_\hbar T_{n-l}(F_{l+1} \otimes \cdots \otimes F_n) + \\
&\sum_{t=1}^{n-1} \sum_{l < k_1 < \cdots < k_t} \sum_{l < l_1 < \cdots < l_{n-t}} \left(\frac{\hbar}{i}\right)^{t-1} T_l(F_1 \otimes \cdots \otimes F_l) \star_\hbar T_{n-l-t+1}(\delta_t(F_{k_1} \otimes \dots F_{k_t}) \otimes F_{l_1} \otimes \dots F_{l_{n-l-t}}) \,.
\end{aligned} \tag{348}$$

We now apply the inductive hypothesis that eq. (346) holds at order $n-1$, together with the fact that s_0 is a graded derivation of $\mathbf{W}_0$ (we proved this above for $\hat{s}_0$, the proof for s_0 is completely analogous). If this is done, then it follows that $M_n(F_1 \otimes \cdots \otimes F_n) = 0$ under the assumed support properties for the F_i. Consequently, M_n must be a functional valued in $\mathbf{W}_0$ that is supported on the total diagonal. That functional must hence be of the form $(\hbar/i)^{n-1} T_1(\delta_n(F_1 \otimes \cdots \otimes F_n))$ for some δ_n, which we hence take as the definition of δ_n.

We must next show that $\delta_n(F^{\otimes n})$ is of order $\hbar$. For this, we pick a quasifree state ω of $\mathbf{W}_0$, and we define, as described in Appendix B, the "connected time ordered products" T^c_ω by the formula

$$T^c_{n,\omega}(F_1 \otimes \cdots \otimes F_n) := T_n(F_1 \otimes \cdots \otimes F_n) - \sum_P :\prod_{J \in P} T_{|J|}\big(\otimes_{j \in J} F_j\big):_\omega \tag{349}$$

where P runs over all partitions of $\{1, \dots, n\}$, and where J runs through the disjoint sets in the given partition. A generating type functional formula can be obtained using the linked cluster theorem, and is given by eq. (486). The key fact about the connected products is that the n-th product is of order $O(\hbar^{n-1})$ if the F_i themselves are of order $O(1)$. This will now be used by formulating eq. (346) in terms of connected products. Using generating functional expression for the connected time ordered products, and using the fact that s_0 is a derivation with respect

to the Wick product (which follows from eq. (298)), one can easily see that

$$
\begin{aligned}
&\left(\frac{i}{\hbar}\right)^{n-1} s_0 T^c_{n,\omega}(F_1\otimes\cdots\otimes F_n) - \sum_{k=0}^{n}\left(\frac{i}{\hbar}\right)^{n-1} T^c_{n,\omega}(F_1\otimes\ldots s_0F_k\otimes\ldots F_n) - \\
&\sum_{t=1}^{n-1}\sum_{k_1<\cdots<k_t}\sum_{l_1<\cdots<l_{n-t}}\left(\frac{i}{\hbar}\right)^{n-t} T^c_{n-t+1,\omega}(\delta_t(F_{k_1}\otimes\ldots F_{k_t})\otimes F_{l_1}\otimes\ldots F_{l_{n-t}}) \\
&= T_1\big(\delta_n(F_1\otimes\cdots\otimes F_n)\big)\,.
\end{aligned} \tag{350}
$$

Now, if we inductively assume that δ_t is of order $O(\hbar)$ for orders $t<n$, then it follows that the order of the second sum in the above expression is $O(\hbar)$. Furthermore, the first two terms on the left side in the above equation precisely cancel up to a term of order $O(\hbar)$. This follows from the fact that the limit $\lim_\hbar T^c_{n,\omega}/\hbar^{n-1}$ correspond to the "tree diagrams", and there are no anomalies at tree level [40]. Thus, $\delta_n = O(\hbar)$, as we desired to show. ☐

We are now in a position to complete the proof. From eqs. (347) and (345) we get the desired Ward identity (333) with

$$
A(\mathrm{e}_\otimes^F) := \Delta_B(\mathrm{e}_\otimes^F) + \delta(\mathrm{e}_\otimes^F)\,, \quad B = \int_M \frac{\delta_R S_0}{\delta\Phi(x)}\wedge\frac{\delta_L F}{\delta\Phi^\ddagger(x)}\,. \tag{351}
$$

We must finally show that the maps A_n have properties analogous to those of the maps D_n in sec. 3.6, i.e., properties (i)—(vi). The proof is similar as the proof for the D_n outlined there. It is again inductive in nature and is based on the expression

$$
\begin{aligned}
&T_1(A_n(F_1\otimes\cdots\otimes F_n)) = \hat{s}_0 T_n(F_1\otimes\cdots\otimes F_n) - \\
&\sum_{k=0}^{n} T_n(F_1\otimes\ldots\hat{s}_0F_k\otimes\ldots F_n) - \frac{\hbar}{i}\sum_{k<j} T_{n-1}(F_1\otimes\ldots(F_j,F_k)\otimes\ldots F_n) - \\
&\sum_{t=1}^{n-1}\sum_{k_1<\cdots<k_t}\sum_{l_1<\cdots<l_{n-t}}\left(\frac{\hbar}{i}\right)^{t-1} T_{n-t+1}(A_t(F_{k_1}\otimes\ldots F_{k_t})\otimes F_{l_1}\otimes\ldots F_{l_{n-t}})
\end{aligned} \tag{352}
$$

for the n-th order anomaly. We have already shown that $A_n = O(\hbar)$, because this is true for Δ_n, δ_n to all orders. The statement (ii) follows because all quantities on the right side of this equation are locally and covariantly constructed out of the metric. (iii) follows from the fact that $\hat{s}_0$ increases the ghost number by 1 unit, and because the anti-bracket increases the ghost number by 1 unit. (iv) follows because $\hat{s}_0$ and the anti-bracket preserve the dimension, and from the known scaling behavior of the time-ordered products, T2. (v) follows because $\hat{s}_0$ is compatible with the *-operation and because the time ordered products are unitary, see T7. For more details on such kinds of arguments, see again [62].

To complete the proof, we must still show that Proposition 4 is indeed true. These arguments are given in detail in thm. 7 and lemma 11 of [15]. For completeness, we here outline a slightly modified version of these arguments, but we refer the reader to this work for full details[14].

[14]The arguments in [15] are given only for the case of flat spacetime, but the key steps easily generalize to curved manifolds straightforwardly.

Proof of Proposition 4: We begin by writing down the n-th order part of eq. (342), given by

$$-T_1\big(\Delta_n(F_1\otimes\cdots\otimes F_n)\big) = \Big(\frac{i}{\hbar}\Big)^n \int T_{n+1}\Big(F_1\otimes\cdots\otimes F_n\otimes\psi(x)\wedge\frac{\delta S_0}{\delta\Phi(x)}\Big) +$$
$$\Big(\frac{i}{\hbar}\Big)^{n-1}\sum_{i=1}^{n}\int T_n\Big(F_1\otimes\ldots\psi(x)\wedge\frac{\delta F_i}{\delta\Phi(x)}\otimes\ldots F_n\Big) -$$
$$\Big(\frac{i}{\hbar}\Big)^n\int T_n\big(F_1\otimes\cdots\otimes F_n\otimes\psi(x)\big)\star_\hbar\frac{\delta S_0}{\delta\Phi(x)} +$$
$$\sum_{t=1}^{n-1}\Big(\frac{i}{\hbar}\Big)^{n-t}\sum_{k_1<\ldots<k_t}\sum_{l_1<\ldots<l_{n-t}} T_{n-t+1}\big(\Delta_t(F_{k_1}\otimes\cdots\otimes F_{k_l})\otimes F_{l_1}\cdots\otimes F_{l_{n-t}}\big)\,. \tag{353}$$

For $n=0$, the identity becomes

$$-T_1(\Delta_0) = \int T_1\big(\psi(x)\wedge\frac{\delta S_0}{\delta\Phi(x)}\big) - \int T_1\big(\psi(x)\big)\star_\hbar\frac{\delta S_0}{\delta\Phi(x)}\,. \tag{354}$$

The function Δ_0 is trivially local in this case. Because the first time ordered product T_1 as well as the $\star_\hbar$-product reduce to the ordinary product in the space of classical local functionals of the fields when $\hbar\to 0$, it follows that $\Delta_0 = O(\hbar)$, as claimed.

We now proceed iteratively in n. We assume that the assertion about Δ_n in the proposition has already been proved for Δ_k up to $k=n-1$. In fact, let us assume for simplicity even that $\Delta_k=0$ up to $k=n-1$. We define $M_n(F_1\otimes\cdots\otimes F_n)$ to be the right side of eq. (353). The aim is to prove that this is a local functional valued in $\mathbf{W}_0$. To demonstrate this, consider functionals F_i with the property that

$$\left(\bigcup_{i=1}^{l}\operatorname{supp}\Big(\frac{\delta F_i}{\delta\Phi}\Big)\cup\operatorname{supp}\psi\right)\cap J^+\left(\bigcup_{i=l+1}^{n}\operatorname{supp}\Big(\frac{\delta F_i}{\delta\Phi}\Big)\right)=\emptyset \tag{355}$$

for some l not equal to n. Then M_n can be written as follows using the causal factorization properties of the time ordered products:

$$M_n\big(F_1\otimes\cdots\otimes F_n\big) = \Big(\frac{i}{\hbar}\Big)^n\int T_{l+1}\Big(F_1\otimes\cdots\otimes F_l\otimes\psi(x)\wedge\frac{\delta S_0}{\delta\Phi(x)}\Big)\star_\hbar T_{n-l}\Big(F_{l+1}\otimes\cdots\otimes F_n\Big) +$$
$$\Big(\frac{i}{\hbar}\Big)^{n-1}\sum_{i=1}^{l}\int T_l\Big(F_1\otimes\ldots\psi(x)\wedge\frac{\delta F_i}{\delta\Phi(x)}\otimes\ldots F_l\Big)\star_\hbar T_{n-l}\Big(F_{l+1}\otimes\cdots\otimes F_n\Big) -$$
$$\Big(\frac{i}{\hbar}\Big)^n\int T_n\big(F_1\otimes\cdots\otimes F_l\otimes\psi(x)\big)\star_\hbar\frac{\delta S_0}{\delta\Phi(x)}\star_\hbar T_{n-l}\Big(F_{l+1}\otimes\cdots\otimes F_n\Big)\,, \tag{356}$$

where we have used that, for any $G\in\mathbf{W}_0$ (of even Grassmann parity), we have the identity

$$G\star_\hbar\frac{\delta S_0}{\delta\Phi(x)} = \frac{\delta S_0}{\delta\Phi(x)}\star_\hbar G\,, \tag{357}$$

which in turn follows from the definition of the star-product given above in sec. 4.2, eq. (300), together with the fact that $\delta S_0/\delta\Phi^i(x) = D_{ij}\Phi^j(x)$, $D_{ij}\omega^{jk}(x,y) = 0$, with the D_{ij} the matrix of linear partial differential operators in the field equation for the free underlying (gauge fixed) theory with action S_0. Using now the inductive assumption in eq. (356), we conclude that M_n for the F_i with the assumed support properties. It follows from this that M_n can only be supported on the diagonal, which is the desired locality property of $\Delta_n(F_1 \otimes \cdots \otimes F_n)$.

It remains to be seen that $\Delta_n = O(\hbar)$. For this, we take eq. (342) and multiply from the left with the anti-time ordered products [see eq. (148)], to obtain

$$\begin{aligned}
&\bar{T}\big(\mathrm{e}_\otimes^{iF/\hbar}\big) \star_\hbar \int T\big(\mathrm{e}_\otimes^{iF/\hbar} \otimes \big[\psi(x) \wedge \frac{\delta F}{\delta\Phi(x)} + \psi(x) \wedge \frac{\delta S_0}{\delta\Phi(x)}\big]\big)\, dx \\
= \quad &\bar{T}\big(\mathrm{e}_\otimes^{iF/\hbar}\big) \star_\hbar \int T\big(\mathrm{e}_\otimes^{iF/\hbar} \otimes \phi(x)\big) \star_\hbar \frac{\delta S_0}{\delta\Phi(x)}\, dx + \bar{T}\big(\mathrm{e}_\otimes^{iF/\hbar}\big) \star_\hbar T\big(\mathrm{e}_\otimes^{iF/\hbar} \otimes \Delta_B(\mathrm{e}_\otimes^F)\big)\,.
\end{aligned} \tag{358}$$

Using next the definition of the retarded products [see eq. (264)], this may be rewritten in the form

$$\begin{aligned}
&\int R\big(\psi(x) \wedge \frac{\delta F}{\delta\Phi(x)} + \psi(x) \wedge \frac{\delta S_0}{\delta\Phi(x)}; \mathrm{e}_\otimes^{iF/\hbar}\big)\, dx \\
= \quad &\int R\big(\psi(x); \mathrm{e}_\otimes^{iF/\hbar}\big) \star_\hbar \frac{\delta S_0}{\delta\Phi(x)}\, dx + R\big(\Delta_B(\mathrm{e}_\otimes^F); \mathrm{e}_\otimes^{iF/\hbar}\big)\,.
\end{aligned} \tag{359}$$

The key point is now the that the retarded products in this equation have a meaningful limit as $\hbar \to 0$, as proven in [37], i.e., the above expressions contain no inverse powers of $\hbar$, despite the inverse powers of $\hbar$ in the exponentials. This limit is just the classical limit for the interacting fields as defined by the Bogoliubov formula eq. (263). Furthermore, the classical limit of $\star_\hbar$ is the usual classical product of classical fields. Thus, the eq. (359) has a classical limit, the "classical Master Ward Identity" of [15]. It is shown in this reference that this identity in classical field theory is indeed true with $\Delta = 0$. Consequently, Δ itself must be of order $O(\hbar)$, as we desired to show. This concludes our outline of the proof of Proposition 4. $\square$.

Since we have proved Proposition 4, we have proved Proposition 3. $\square$

We next derive a “consistency condition” on the anomaly.

Proposition 5 ("Consistency condition") The anomaly satisfies the equation

$$\boxed{\Big(S_0 + F, A(\mathrm{e}_\otimes^F)\Big) - \frac{1}{2}A\Big((S_0 + F, S_0 + F) \otimes \mathrm{e}_\otimes^F\Big) = A\Big(A(\mathrm{e}_\otimes^F) \otimes \mathrm{e}_\otimes^F\Big)\,.} \tag{360}$$

Proof of Proposition 5: We first act with $\hat{s}_0$ on the anomalous Ward identity eq. (333) and use that $\hat{s}_0^2 = 0$. We obtain the equation

$$0 = \hat{s}_0 T\left(A(\mathrm{e}_\otimes^F) \otimes \mathrm{e}_\otimes^{iF/\hbar}\right) + \frac{1}{2}\hat{s}_0 T\left((S_0 + F, S_0 + F) \otimes \mathrm{e}_\otimes^{iF/\hbar}\right) = (\mathrm{I}) + (\mathrm{II}) \tag{361}$$

The trick is now to apply the anomalous Ward identity one more time to each of the terms on the right side. For simplicity, we assume that F has Grassmann parity 0. We can then write the first term as

$$
\begin{aligned}
&(\mathrm{I}) = \frac{\hbar}{i}\frac{d}{d\tau}\hat{s}_0 T\left(\mathrm{e}_\otimes^{i(F+\tau A(\mathrm{e}_\otimes^F))}\right)\Big|_{\tau=0} = \\
&\frac{d}{d\tau}\Bigg[\frac{1}{2}T\left(\left(S_0+F+\tau A(\mathrm{e}_\otimes^F), S_0+F+\tau A(\mathrm{e}_\otimes^F)\right)\otimes \mathrm{e}_\otimes^{i(F+\tau A(\mathrm{e}_\otimes^F))/\hbar}\right) + \\
&T\left(A(\mathrm{e}_\otimes^{\tau A(\mathrm{e}_\otimes^F)})\otimes e^{i(F+\tau A(\mathrm{e}_\otimes^F))/\hbar}\right)\Bigg]\Bigg|_{\tau=0} = \\
&T\left((S_0+F, A(\mathrm{e}_\otimes^F))\otimes \mathrm{e}_\otimes^{iF/\hbar}\right) + \frac{i}{2\hbar}T\left(A(\mathrm{e}_\otimes^F)\otimes(S_0+F,S_0+F)\otimes \mathrm{e}_\otimes^{iF/\hbar}\right) - \\
&T\left(A(A(\mathrm{e}_\otimes^F)\otimes \mathrm{e}_\otimes^F)\otimes \mathrm{e}_\otimes^{iF/\hbar}\right) + T\left(A(\mathrm{e}_\otimes^F)\otimes A(\mathrm{e}_\otimes^F)\otimes \mathrm{e}_\otimes^{iF/\hbar}\right). \qquad (362)
\end{aligned}
$$

Since F has Grassmann parity 0, $A(\mathrm{e}_\otimes^F)$ has Grassmann parity 1, so by the anti-symmetry of the time-ordered products for such elements, see (241), the last term vanishes. Next, we apply the anomalous Ward identity to term (II). We now obtain

$$
\begin{aligned}
&(\mathrm{II}) = \frac{\hbar}{2i}\frac{d}{d\tau}\hat{s}_0 T\left(\mathrm{e}_\otimes^{i(F+\tau(S_0+F,S_0+F))/\hbar}\right)\Big|_{\tau=0} = \\
&\frac{1}{2}\frac{d}{d\tau}\Bigg[\frac{1}{2}T\left((S_0+F+\tau(S_0+F,S_0+F), S_0+F+\tau(S_0+F,S_0+F))\otimes \mathrm{e}_\otimes^{i(F+\tau(S_0+F,S_0+F))/\hbar}\right) + \\
&T\left(A(\mathrm{e}_\otimes^{\tau(S_0+F,S_0+F)})\otimes \mathrm{e}_\otimes^{i(F+\tau(S_0+F,S_0+F))/\hbar}\right)\Bigg]\Bigg|_{\tau=0} = \\
&\frac{1}{2}\Bigg[T\left((S_0+F,(S_0+F,S_0+F))\otimes \mathrm{e}_\otimes^{iF/\hbar}\right) + \\
&\frac{i}{2\hbar}T\left((S_0+F,S_0+F)\otimes(S_0+F,S_0+F)\otimes \mathrm{e}_\otimes^{iF/\hbar}\right) - \\
&T\left(A((S_0+F,S_0+F)\otimes \mathrm{e}_\otimes^F)\otimes \mathrm{e}_\otimes^{iF/\hbar}\right) - \frac{i}{\hbar}T\left(A(\mathrm{e}_\otimes^F)\otimes(S_0+F,S_0+F)\otimes \mathrm{e}_\otimes^{iF/\hbar}\right)\Bigg]. \qquad (363)
\end{aligned}
$$

Now, the first term on the right side vanishes due to the graded Jacobi identity (59) for the anti-bracket. The second term vanishes due to the anti-symmetry property of the time ordered products (241), since (S_0+F,S_0+F) has Grassmann parity 1. If we now add up terms (I) and (II), we end up with the following identity:

$$
\begin{aligned}
&T\left(\left[(S_0+F,A(\mathrm{e}_\otimes^F)) - \frac{1}{2}A((S_0+F,S_0+F)\otimes \mathrm{e}_\otimes^F)\right]\otimes \mathrm{e}_\otimes^{iF/\hbar}\right) = \\
&T\left(A(A(\mathrm{e}_\otimes^F)\otimes \mathrm{e}_\otimes^F)\otimes \mathrm{e}_\otimes^{iF/\hbar}\right). \qquad (364)
\end{aligned}
$$

Since a time ordered product $T(G \otimes e_{\otimes}^{iF/\hbar})$ vanishes if and only if $G = 0$, the desired consistency condition (360) follows $\square$

Let us summarize what we have shown so far: We first demonstrated that the Ward identity (336) always holds with an anomaly term of order $\hbar$, i.e., that eq. (333) holds. We then showed that the anomaly is not arbitrary, but must obey the consistency condition (360). This condition imposes a strong restriction on the possible anomalies, and we will show in the following subsections using this condition that, when F is as in the cases T12a, T12b, and T12c, then the anomaly $A(e_{\otimes}^{F})$ can in fact be removed by a redefinition of the time-ordered products consistent with T1—T11. Thus, in these cases, we may achieve that the Ward identity (336) holds exactly, without anomaly.

To prepare the proof of this statement, we first note that, since the anomaly itself is of order $\hbar$, the lowest order in $\hbar$ contribution to the "anomaly of the anomaly term" on the right side of eq. (360) is necessarily of a higher order in $\hbar$ than the lowest order contribution left side. An even more stringent consistency condition can therefore be obtained for the lowest order (in $\hbar$) contribution to the anomaly. For this, we expand $A(e_{\otimes}^{F})$ in powers of the coupling, λ, and $\hbar$,

$$A(e_{\otimes}^{F}) = \sum_{n,m>0} \hbar^m \frac{\lambda^n}{n!} \int \mathcal{A}_n^m(x_1,\dots,x_n)\, f(x_1)\dots f(x_n)\, dx_1 \dots dx_n\,, \tag{365}$$

where $\mathcal{A}_n^m$ is a local, covariant functional of $(\Phi, \Phi^{\ddagger})$, and the metric that is supported on the total diagonal. Both sums start with positive powers, because the anomaly vanishes in the classical theory (i.e., $\hbar = 0$), and also in the free quantum theory (i.e., $\lambda = 0$). An explicit definition of $\mathcal{A}_n^m$ is given by

$$\mathcal{A}_n^m(x_1,\dots,x_n) = \frac{1}{m!}\frac{\partial^m}{\partial\hbar^m}\frac{\delta^n}{\delta f(x_1)\cdots\delta f(x_n)} A(e_{\otimes}^{F})\Big|_{f=0=\hbar}\,. \tag{366}$$

Let $A^m(e_{\otimes}^{F})$ now be the lowest order contribution to $A(e_{\otimes}^{F})$ *in the $\hbar$-expansion*, that is, m is the smallest integer for which

$$A^m(e_{\otimes}^{F}) := \frac{1}{m!}\frac{\partial^m}{\partial\hbar^m} A(e_{\otimes}^{F})\Big|_{\hbar=0} \tag{367}$$

is not zero. (Note that the quantity A^m is different from the quantity A_n above!) Then, from our consistency condition given in Proposition 5, we get the following version of the consistency condition:

Proposition 6: ("$\hbar$-expanded consistency condition") Let A be the anomaly of the Ward identity in Proposition 3, and let A^m be the first non-trivial term in the $\hbar$-expansion of A. Then we have

$$\boxed{\Big(S_0 + F, A^m\big(e_{\otimes}^{F}\big)\Big) - \frac{1}{2}A^m\Big(\big(S_0 + F, S_0 + F\big) \otimes e_{\otimes}^{F}\Big) = 0\,.} \tag{368}$$

Here, $(\,.\,,\,.\,)$ is the anti-bracket [see eq. (58)].

This stronger form of the consistency condition is the key relation that will be used in the proofs of T12a, T12b, and T12c. In those proofs we will actually encounter several quantities like $\mathcal{A}^m_n$, so it is convenient to use again the notation from sec. 3.6. As there, $(k_1,\dots,k_n)$ is a set of natural numbers. We denote by $\mathbf{P}^{k_1/\dots/k_n}(M^n)$ the space of all local, covariant functionals of $\Phi,\Phi^\ddagger$, and the metric which are supported on the total diagonal, and which take values in the bundle (248) of antisymmetric tensors over M^n. Thus, if $\mathcal{B}_n \in \mathbf{P}^{k_1/\dots/k_n}(M^n)$, then $\mathcal{B}_n$ is a (distributional) polynomial, local, covariant functional of $\Phi,\Phi^\ddagger$ and the metric taking values in the $k_1+\dots+k_n$ forms over M^n, which is supported on the total diagonal. It is a k_1-form in the first variable x_1, a k_2-form in the second variable x_2, etc. Concerning such quantities, we have a simple lemma that we will use below.

Lemma 9. Let $\mathcal{B}_n \in \mathbf{P}^{k_1/\dots/k_n}(M^n)$, and let $f_i, i=1,\dots,n$ be closed forms on M of degree $4-k_i$. Assume that for any such forms, we have

$$\int \mathcal{B}_n(x_1,\dots,x_n)\wedge\prod_i f_i(x_i) = 0\,. \tag{369}$$

Then it is possible to write

$$\mathcal{B}_n[\Phi,\Phi^\ddagger] = \sum_{k=1}^n d_k\mathcal{B}_{n/k}[\Phi,\Phi^\ddagger] + \mathcal{B}_n[0,0]\,, \tag{370}$$

where $d_k = dx_k\wedge(\partial/\partial x^\mu_k)$ is the exterior differential applied to the k-th variable.

Proof: We first consider the case $n=1$. If $k_1=4$, then the assumptions imply that $F = \int\mathcal{B}_1(x)\,f_1(x) = 0$ for any closed 0-form f_1, i.e., for any constant such as $f_1(x)=1$. We therefore have $\delta F/\delta\psi(x) = 0$, using the abbreviation $\psi = (\Phi,\Phi^\ddagger)$. Consider the path $\psi_\tau = (\tau\Phi,\tau\Phi^\ddagger)$ in field space. Then

$$\frac{d}{d\tau}\mathcal{B}_1[\psi_\tau] = \sum_k(\nabla^k\psi)\frac{\partial\mathcal{B}_1[\psi_\tau]}{\partial(\nabla^k\psi)} = \psi\frac{\delta F[\psi_\tau]}{\delta\psi} + d\vartheta[\psi_\tau] = d\vartheta[\psi_\tau]\,, \tag{371}$$

for some locally constructed 3-form ϑ. Thus,

$$\begin{aligned}\mathcal{B}_1[\Phi,\Phi^\ddagger] &= \mathcal{B}_1[0,0] + \int_0^1\frac{d}{d\tau}\mathcal{B}_1[\psi_\tau]\,d\tau = \mathcal{B}_1[0,0] + d\int_0^1\vartheta[\psi_\tau]\,d\tau && (372)\\ &= \mathcal{B}_1[0,0] + d\mathcal{B}_{1/1}[\Phi,\Phi^\ddagger]\,, && (373)\end{aligned}$$

which has the desired form. If $k_1=0$, then f_1 is a 4-form, which is always closed. Thus, the assumptions of the lemma imply that $\mathcal{B}_1[\Phi,\Phi^\ddagger]=0$, which is again of the desired form. Finally, if $0<k_1<4$, we may choose $f_1 = dh_1$, implying that $\int d\mathcal{B}_1(x)\wedge h_1(x) = 0$ for all h_1, and thus that $d\mathcal{B}_1 = 0$. The statement now follows from the algebraic Poincare lemma.

The proof of the lemma for $n > 1$ can now be generalized from the case $n = 1$. Without loss of generality, we may assume $\mathcal{B}_n[\Phi = 0, \Phi^\ddagger = 0] = 0$, for otherwise, we may simply subtract this quantity. To reduce the situation to $n = 1$, consider the form on M of degree k_1 that is obtained by smearing $\mathcal{B}_n$ as in (374), but the smearing over the first test-form f_1 omitted. If $k_1 < 4$, then this form is a closed form that is locally and covariantly constructed from $f_2, f_2, \dots, f_n$ and $\Phi, \Phi^\ddagger$. This 4-form then by definition obeys the assumptions of lemma 3, so we may write

$$0 = \int \mathcal{B}_n(x_1, \dots, x_n) \wedge \prod_{i=2}^n f_i(x_i) - d_1 \int \mathcal{B}_{n/1}(x_1, \dots, x_n) \wedge \prod_{i=2}^n f_i(x_i) \tag{374}$$

for some $\mathcal{B}_{n/1} \in \mathbf{P}^{k_1-1/k_2/\dots/k_n}$. If $k_1 = 4$ one may argue similarly. We now repeat this argument, now omitting the integration over the second test form f_2. We then get

$$\begin{aligned} 0 = \int \mathcal{B}_n(x_1, \dots, x_n) \wedge \prod_{i=3}^n f_i(x_i) \\ - d_1 \int \mathcal{B}_{n/1}(x_1, \dots, x_n) \wedge \prod_{i=3}^n f_i(x_i) - d_2 \int \mathcal{B}_{n/2}(x_1, \dots, x_n) \wedge \prod_{i=3}^n f_i(x_i) \end{aligned} \tag{375}$$

for some $\mathcal{B}_{n/2} \in \mathbf{P}^{k_1/k_2-1/\dots/k_n}$. We may continue this procedure, and thus inductively proceed to construct the remaining $\mathcal{B}_{n/k}$. □

4.4.1 Proof of T12a

Up to now, we have shown (Proposition 3) that any prescription for defining time ordered products satisfying properties T1-T11 satisfies the Ward identity (333) with anomaly. We shall now prove that we can change the definition of the time ordered products in such a way that T1-T11 still hold, and such that in addition the anomaly vanishes in the case when $F = \int \{\lambda f \mathbf{L}_1 + \lambda^2 f^2 \mathbf{L}_2\}$, where $f \in C_0^\infty(M)$. Thus, our new prescription will satisfy (322) [and in fact even eq. (336)] for this F. This will then enable us to prove that the new prescription for defining time ordered products will satisfy property T12a.

The key tool for proving this statement is the consistency condition on the $\hbar$-expanded anomaly given in Proposition 6. To take full advantage of this consistency condition, we would like to put $f = 1$, for we then have $S_0 + F = S$, and we can take advantage of BRST-invariance of the full action S, see (57). We note that we cannot simply set $f = 1$ in $T(\mathrm{e}_\otimes^{iF/\hbar})$, for we might encounter infra-red divergences. However, since the anomaly terms $\mathcal{A}_n^m$ are local, covariant functionals of $\Phi, \Phi^\ddagger$ that are supported on the total diagonal (taking values in the $4n$-forms $\wedge^{4n} T^* M^n$ over M^n), we may without any danger set $f = 1$ in eq. (368). As we have already said, in that case we have $F = \lambda S_1 + \lambda^2 S_2$, and consequently $S_0 + F = S$, where S is the full action (57). So from eq. (368) together with $(S, S) = 0$ and $\hat{s} = (S, \,.\,)$ we find

$$\hat{s} A^m \big(\mathrm{e}_\otimes^{\lambda S_1 + \lambda^2 S_2}\big) = 0 \,. \tag{376}$$

Now, we have

$$A^m\big(\mathrm{e}_\otimes^{\lambda S_1+\lambda^2 S_2}\big) = \int_M a^m(x) = \sum_{n>0} \frac{\lambda^n}{n!} \int_M a_n^m(x)\,, \tag{377}$$

where $a^m \in \mathbf{P}^4(M)$ (and likewise for a_n^m). Furthermore, from the properties of the anomaly derived in the previous subsection, the dimension of a^m must be 4, and the ghost number must be $+1$. Equation (376) may now be viewed as saying that $a^m \in H^1(\hat{s}|d,\mathbf{P}^4)$. From the Lemmas given in sec. 2.2, we have a complete classification of all the elements in this ring. In fact, as shown there in lemma 1, all non-trivial elements in this ring at ghost number +1 and dimension 4 must be even under parity, $\varepsilon \to -\varepsilon$ when the Lie-group has no abelian factors. On the other hand, it follows from the properties of the anomaly A that a^m is parity odd, i.e., $a^m \to -a^m$ under parity $\varepsilon \to -\varepsilon$. Therefore, a^m must represent the zero element in the ring $H^1(\hat{s}|d,\mathbf{P}^4)$, so there are $b^m \in \mathbf{P}^4_0(M)$ and $c^m \in \mathbf{P}^3_1(M)$ such that

$$a^m(x) = \hat{s}b^m(x) + dc^m(x)\,. \tag{378}$$

We expand

$$b_m(x) = \sum_{n>0} \frac{\lambda^n}{n!} b_n^m(x)\,. \tag{379}$$

We would like to use the coefficients $b_n^m(x)$ to redefine the time ordered products $T_n(\mathbf{L}_1(x_1)\otimes \dots \mathbf{L}_1(x_n))$ containing n factors of the interaction Lagrangian. Recalling that by thm. 2, the changes in the time-ordered products are parametrized by local, covariant maps $D_n : \mathbf{P}^{p_1}(M)\otimes \cdots \otimes \mathbf{P}^{p_n}(M) \to \mathbf{P}^{p_1/\dots/p_n}(M^n)$, we define

$$D_n(\mathbf{L}_1(x_1)\otimes\cdots\otimes\mathbf{L}_1(x_n)) := -\hbar^m b_m^n(x_1)\delta(x_1,\dots,x_n)\,. \tag{380}$$

It can be shown that this is within the allowed renormalization freedom for the time-ordered products described in sec. 3.6: First, the locality and covariance of D_n follows from the corresponding property of b_n^m. The scaling property (254) follows from the fact that b_n^m has dimension 4, together with the scaling degree property $sd\,\delta = 4(n-1)$ for the delta function of n spacetime arguments concentrated on the diagonal in M^n. The smooth and analytic dependence of D_n under changes of the spacetime metric again follows from the corresponding properties of b_n^m, while the symmetry is manifest. The unitarity condition (255) follows from the fact that b_n^m is real, which in turn follows from the corresponding property of the anomaly A derived in the previous subsection. To satisfy the field independence property (253), it is furthermore necessary to also change the time-ordered products of sub-monomials of $\mathbf{L}_1$ in order to be consistent with T9. This causes no problems. The identity (256) can be satisfied by defining D_n appropriately for entries o_i that are exterior differentials of $\mathbf{L}_1$. This does not lead to any potential consistency problems, because $\mathbf{L}_1$ itself is not the exterior differential of a locally constructed 3-form. For details of such kinds of arguments see [66], where a very similar situation was treated. Thus, the above D_n (together with the corresponding D_n for sub-Wick monomials of $\mathbf{L}_1$ and their exterior derivatives) gives a permissible change in the time ordered products, i.e.,

the changed time-ordered products $\hat{T}$ defined according to (251) with the above D again satisfy T1–T11.

With the above definition of D_n, and $F = \int[\lambda f \mathbf{L}_1 + \lambda^2 f^2 \mathbf{L}_2]$, the relevant newly defined time-ordered products $\hat{T}$ are given by [see eq. (251)]

$$\begin{aligned}
\hat{T}\left(\mathrm{e}_\otimes^{iF/\hbar}\right) &= T\left(\mathrm{e}_\otimes^{i[F+D(\exp_\otimes F)]/\hbar}\right) && (381)\\
\hat{T}\left((S_0+F,S_0+F)\otimes \mathrm{e}_\otimes^{iF/\hbar}\right) &= T\left((S_0+F,S_0+F)\otimes \mathrm{e}_\otimes^{i[F+D(\exp_\otimes F)]/\hbar}\right) && (382)\\
\hat{T}\left(\hat{A}(\mathrm{e}_\otimes^F)\otimes \mathrm{e}_\otimes^{iF/\hbar}\right) &= T\left([\hat{A}(\mathrm{e}_\otimes^F)+D(\hat{A}(\mathrm{e}_\otimes^F)\otimes \mathrm{e}_\otimes^F)]\otimes \mathrm{e}_\otimes^{i[F+D(\exp_\otimes F)]/\hbar}\right), && (383)
\end{aligned}$$

where $\hat{A}(\mathrm{e}_\otimes^F)$ is the anomaly (333) in the Ward identity for the modified time ordered products $\hat{T}$. The second line follows because there are no sub-Wick monomials in $\mathbf{L}_1$ that are also contained in (S_0+F,S_0+F), as one may check explicitly. We would now like to relate the anomaly $A(\mathrm{e}_\otimes^F)$ of the "old" time ordered products T to the anomaly $\hat{A}(\mathrm{e}_\otimes^F)$ of the "new" time ordered products $\hat{T}$. We have

$$\begin{aligned}
&\frac{i}{\hbar}T\left([\hat{A}(\mathrm{e}_\otimes^F)+D(\hat{A}(\mathrm{e}_\otimes^F)\otimes \mathrm{e}_\otimes^F)]\otimes \mathrm{e}_\otimes^{i[F+D(\exp_\otimes F)]/\hbar}\right)\\
=\;& \frac{i}{\hbar}\hat{T}\left(\hat{A}(\mathrm{e}_\otimes^F)\otimes \mathrm{e}_\otimes^{iF/\hbar}\right)\\
=\;& \hat{s}_0\hat{T}\left(\mathrm{e}_\otimes^{iF/\hbar}\right)-\frac{i}{2\hbar}\hat{T}\left((S_0+F,S_0+F)\otimes \mathrm{e}_\otimes^{iF/\hbar}\right)\\
=\;& \hat{s}_0 T\left(\mathrm{e}_\otimes^{i[F+D(\exp_\otimes F)]/\hbar}\right)-\frac{i}{2\hbar}T\left((S_0+F,S_0+F)\otimes \mathrm{e}_\otimes^{i[F+D(\exp_\otimes F)]/\hbar}\right)\\
=\;& \frac{i}{2\hbar}T\left((S_0+F+D(\mathrm{e}_\otimes^F),S_0+F+D(\mathrm{e}_\otimes^F))\otimes \mathrm{e}_\otimes^{i[F+D(\exp_\otimes F)]/\hbar}\right)\\
-\;& \frac{i}{2\hbar}T\left((S_0+F,S_0+F)\otimes \mathrm{e}_\otimes^{i[F+D(\exp_\otimes F)]/\hbar}\right)\\
+\;& \frac{i}{\hbar}T\left(A(\mathrm{e}_\otimes^{F+D(\exp_\otimes F)})\otimes \mathrm{e}_\otimes^{i[F+D(\exp_\otimes F)]/\hbar}\right)\\
=\;& \frac{i}{\hbar}T\left((S_0+F,D(\mathrm{e}_\otimes^F))\otimes \mathrm{e}_\otimes^{i[F+D(\exp_\otimes F)]/\hbar}\right)+\frac{i}{\hbar}T\left(A(\mathrm{e}_\otimes^{F+D(\exp_\otimes F)})\otimes \mathrm{e}_\otimes^{i[F+D(\exp_\otimes F)]/\hbar}\right)\\
+\;& \frac{i}{2\hbar}T\left((D(\mathrm{e}_\otimes^F),D(\mathrm{e}_\otimes^F))\otimes \mathrm{e}_\otimes^{i[F+D(\exp_\otimes F)]/\hbar}\right). \qquad (384)
\end{aligned}$$

It follows from this equation that

$$\hat{A}\big(\mathrm{e}_\otimes^F\big)+D\big(\hat{A}(\mathrm{e}_\otimes^F)\otimes \mathrm{e}_\otimes^F\big)=\big(S_0+F,D(\mathrm{e}_\otimes^F)\big)+A\big(\mathrm{e}_\otimes^{F+D(\exp_\otimes F)}\big)+\frac{1}{2}\big(D(\mathrm{e}_\otimes^F),D(\mathrm{e}_\otimes^F)\big). \qquad (385)$$

We now use that, by definition, $D(\mathrm{e}_\otimes^F)=O(\hbar^m)$, and by assumption $A(\mathrm{e}_\otimes^F)=O(\hbar^m)=\hat{A}(\mathrm{e}_\otimes^F)$. Then it follows that, at order $\hbar^m$, we have

$$\hat{A}^m\big(\mathrm{e}_\otimes^F\big)=A^m\big(\mathrm{e}_\otimes^F\big)+\big(S_0+F,D^m(\mathrm{e}_\otimes^F)\big). \qquad (386)$$

Now it follows from our definition of D and eq. (378) that $\hat{s}D^m(\mathrm{e}_\otimes^{\lambda S_1+\lambda^2 S_2}) = -A^m(\mathrm{e}_\otimes^{\lambda S_1+\lambda^2 S_2})$. Therefore, if we now put $f = 1$ in $F = \int[\lambda f \mathbf{L}_1 + \lambda^2 f^2 \mathbf{L}_2]$ in the above equation, then we find

$$\hat{A}^m\big(\mathrm{e}_\otimes^{\lambda S_1+\lambda^2 S_2}\big) = A^m\big(\mathrm{e}_\otimes^{\lambda S_1+\lambda^2 S_2}\big) + \hat{s}D^m\big(\mathrm{e}_\otimes^{\lambda S_1+\lambda^2 S_2}\big) = 0\,. \tag{387}$$

Thus, by our redefinition of the time ordered products, we have already removed the anomaly for any constant test function f. We will now use this fact to completely remove the anomaly by a further redefinition of the time ordered products.

To simplify the notation, we will now again use the notations T and A for the redefined time ordered products and new anomaly, instead of $\hat{T}$ and $\hat{A}$. The anomaly may be expanded in powers of $\hbar$ and λ as in eq. (365). From eq. (387) (remembering that A now denotes $\hat{A}$), we then have

$$\int \mathcal{A}_n^m(x_1,\dots,x_n)\,dx_1\dots dx_n = 0\,, \tag{388}$$

because we can assume at this stage that the anomaly vanishes for constant f. Consequently, by lemma 9, this quantity must be given by an expression of the form

$$\mathcal{A}_n^m(x_1,\dots,x_n) = \sum_{k=1}^{n} d_k \mathcal{C}_{n/k}^m(x_1,\dots,x_n)\,, \tag{389}$$

for some $\mathcal{C}_{n/k}^m \in \mathbf{P}^{4/\dots 3/\dots/4}(M^n)$. We next define a set of D_n by the formula

$$D_n(\mathbf{L}_1(x_1)\otimes\dots\mathcal{O}_1(x_k)\otimes\dots\mathbf{L}_1(x_n)) := -\hbar^m \mathcal{C}_{n/k}^m(x_1,\dots,x_n)\,, \tag{390}$$

where $\mathcal{O}_1 \in \mathbf{P}_1^3(M)$ is the field determined by the equation

$$\hat{s}_0 \mathbf{L}_1 = d\mathcal{O}_1\,, \tag{391}$$

and is given explicitly by

$$\mathcal{O}_1 := f_{IJK} C^I A^J \wedge *dA^K + \frac{1}{2} f_{IJK} C^I C^K * d\bar{C}^K\,. \tag{392}$$

We may again argue that this D_n satisfies all the required properties for an allowed redefinition of the time ordered products, and we denote the new time ordered products again by $\hat{T}$, and the new anomaly again by $\hat{A}$. By a calculation similar to the one given above, the new anomaly now satisfies

$$\hat{A}\big(\mathrm{e}_\otimes^F\big) = A\big(\mathrm{e}_\otimes^F\big) + D\big(A(\mathrm{e}_\otimes^F)\otimes\mathrm{e}_\otimes^F\big) + \frac{1}{2}D\big((S_0+F, S_0+F)\otimes\mathrm{e}_\otimes^F\big)\,. \tag{393}$$

Again evaluating this new anomaly at order $\hbar^m$, we find

$$\hat{A}^m\big(\mathrm{e}_\otimes^F\big) = A^m\big(\mathrm{e}_\otimes^F\big) + \frac{1}{2}D^m\big((S_0+F, S_0+F)\otimes\mathrm{e}_\otimes^F\big)\,. \tag{394}$$

However, our D_n are designed precisely in such a way that $D^m((F,F)\otimes \mathrm{e}_\otimes^F)=0$ and that $D^m(\hat{s}_0 F\otimes \mathrm{e}_\otimes^F)=-A^m(\mathrm{e}_\otimes^F)$, so we find that $\hat{A}^m(\mathrm{e}_\otimes^F)=0$.

In summary, our subsequent definitions of the time ordered products remove the anomaly $A^m(\mathrm{e}_\otimes^F)$ at order $\hbar^m$, and to all orders in λ. We now repeat the same argument for $A^{m+1}(\mathrm{e}_\otimes^F)$, i.e., order $\hbar^{m+1}$, and we can proceed in just the same way for any order in $\hbar$. This shows that the anomaly can be removed to arbitrary orders in $\hbar$ and λ by a redefinition of the time ordered products that is compatible with T1–T11. The absence of an anomaly in eq. (333) for our choice of F implies that T12a is satisfied, because eq. (333) is a generating identity of the identities in T12a. $\square$

4.4.2 Proof of T12b

The proof that the time ordered products can be adjusted, if necessary, so that T12b is satisfied is very similar in nature as that given above for T12a. We therefore only focus on the essential differences.

Consider the local elements $G=\int \gamma\wedge(\mathbf{J}_0+f\lambda\mathbf{J}_1)$ and $F=\int(f\lambda\mathbf{L}_1+f^2\lambda^2\mathbf{L}_2)$, where γ is a smooth 1-form of compact support, and f is a smooth scalar function of compact support. The satisfaction of T12b means that the anomaly in

$$\begin{aligned}\hat{s}_0 T\big(G\otimes \mathrm{e}_\otimes^{iF/\hbar}\big) &= T\big((S_0+F,G)\otimes \mathrm{e}_\otimes^{iF/\hbar}\big)+\frac{i}{2\hbar}T\big((S_0+F,S_0+F)\otimes G\otimes \mathrm{e}_\otimes^{iF/\hbar}\big)\\ &+\ T\big(A(G\otimes \mathrm{e}_\otimes^F)\otimes \mathrm{e}_\otimes^{iF/\hbar}\big)+T\big(A(\mathrm{e}_\otimes^F)\otimes G\otimes \mathrm{e}_\otimes^{iF/\hbar}\big)\end{aligned}\tag{395}$$

can be removed by a suitable redefinition of the time-ordered products. As above, we write

$$A(G\otimes \mathrm{e}_\otimes^F)=\sum_{m,n>0}\hbar^m\frac{\lambda^n}{n!}\int_{M^n}\mathcal{A}_{m,n}(x_1,\dots,x_n)\gamma(x_1)f(x_1)\dots f(x_n)\,dx_1\dots dx_n\,. \tag{396}$$

Let $A^m(G\otimes \mathrm{e}_\otimes^F)$ be the lowest order contribution in $\hbar$ to the anomaly. Because the anomaly is of order at least $\hbar$, we have $m>0$. We apply the consistency condition (368) to the element $F+\tau G$ instead of F in that formula, and we differentiate with respect to τ and set $\tau=0$. Then we obtain the consistency condition

$$\big(S_0+F,A^m(G\otimes \mathrm{e}_\otimes^F)\big)+A^m\big((S_0+F,G)\otimes \mathrm{e}_\otimes^F\big)-\frac{1}{2}A^m\big((S_0+F,S_0+F)\otimes G\otimes \mathrm{e}_\otimes^F\big)=0\,. \tag{397}$$

Now, we put $f=1$ and we take γ to satisfy $d\gamma=0$. Then, $F=\lambda S_1+\lambda^2 S_2$, and $S_0+F=S$, where S is the full action (57) satisfying $(S,S)=0$. Furthermore, by $\hat{s}\mathbf{J}=d\mathbf{K}$,

$$(S_0+F,G)=(S,G)=\hat{s}\int_M\gamma\wedge\mathbf{J}=\int_M\gamma\wedge d\mathbf{K}=-\int_M d\gamma\wedge\mathbf{K}=0\,. \tag{398}$$

Thus, condition (397) implies the condition

$$\hat{s}A^m\big(G\otimes \mathrm{e}_\otimes^{\lambda S_1+\lambda^2 S_2}\big)=0 \tag{399}$$

when γ is closed. Now, we have

$$A^m\big(G\otimes e_\otimes^{\lambda S_1+\lambda^2 S_2}\big) = \int_M \gamma\wedge h^m(x) = \sum_{n>0}\frac{\lambda^n}{n!}\int_M \gamma\wedge h^m_n(x)\,, \tag{400}$$

where $h^m \in \mathbf{P}^3(M)$ (and likewise for h^m_n). Furthermore, from the properties of the anomaly derived in the previous subsection, the dimension of h^m must be 3, and the ghost number must be $+2$. Equation (399), which holds for all closed 1-forms γ in the definition of G, may now be viewed as saying that $h^m \in H^2(\hat{s}|d,\mathbf{P}^3)$. From the Lemmas given in sec. 2.2, we again have a complete classification of all the elements in this ring. In fact, as shown there in lemma 1, all non-trivial elements in this ring at ghost number +2 and dimension 3 must be even under parity, $\varepsilon\to-\varepsilon$ when the Lie-group has no abelian factors. On the other hand, it follows again from the properties of the anomaly A that h^m is parity odd, i.e., $h^m\to -h^m$ under parity $\varepsilon\to-\varepsilon$. Therefore, h^m must represent the zero element in the ring $H^2(\hat{s}|d,\mathbf{P}^3)$, so there are $j^m\in\mathbf{P}^3_1(M)$ and $k^m\in\mathbf{P}^2_2(M)$ such that

$$h^m(x) = \hat{s}j^m(x) + dk^m(x)\,. \tag{401}$$

We again expand j^m in powers of λ

$$j^m(x) = \sum_{n>0}\frac{\lambda^n}{n!}j^m_n(x)\,. \tag{402}$$

Similar to the proof of T12a, we would like to use the coefficients $j^m_n(x)$ to redefine the time ordered products $T_n(\mathbf{J}_1(x_1)\otimes\mathbf{L}_1(x_2)\otimes\dots\mathbf{L}_1(x_n))$ containing $n-1$ factors of the interaction Lagrangian and one factor of the free BRST-current. By thm. 2, the changes in the time-ordered products are parametrized by local, covariant maps $D_n : P^{p_1}(M)\otimes\cdots\otimes P^{p_n}(M)\to\mathbf{P}^{p_1/\cdots/p_n}(M^n)$, and we define

$$D_n(\mathbf{J}_1(x_1)\otimes\mathbf{L}_1(x_2)\otimes\cdots\otimes\mathbf{L}_1(x_n)) := -\hbar^m j^m_n(x_1)\delta(x_1,\dots,x_n)\,. \tag{403}$$

This gives changed time-ordered products via (251), and one may argue as above in T12a that these again satisfy T1–T11. By T12a, we may assume that $A(e_\otimes^F)=0$. The above changes in the time ordered products effected by the maps D_n do not invalidate this, i.e., the new time ordered products $\hat{T}$ will also satisfy $\hat{A}(e_\otimes^F)$, where $\hat{A}$ is the anomaly in eq. (333) for the new time ordered products $\hat{T}$. From eq. (333), we then have, for $F=\int[\lambda f\mathbf{L}_1+\lambda^2 f^2\mathbf{L}_2]$ and any G,

$$\begin{aligned}\hat{s}_0 T\big(G\otimes e_\otimes^{iF/\hbar}\big) &= T\big((S_0+F,G)\otimes e_\otimes^{iF/\hbar}\big)\\ &+\frac{i}{2\hbar}T\big((S_0+F,S_0+F)\otimes G\otimes e_\otimes^{iF/\hbar}\big)+T\big(A(G\otimes e_\otimes^F)\otimes e_\otimes^{iF/\hbar}\big)\end{aligned} \tag{404}$$

and likewise for the "hattet" time ordered products and the hatted anomaly. From this, and the

definition of D above, one finds, for F, G as above

$$\begin{aligned}
& T\big(D(\hat{A}(G \otimes \mathrm{e}_{\otimes}^{F}) \otimes \mathrm{e}_{\otimes}^{F}) \otimes \mathrm{e}_{\otimes}^{iF/\hbar}\big) + T\big(\hat{A}(G \otimes \mathrm{e}_{\otimes}^{F}) \otimes \mathrm{e}_{\otimes}^{iF/\hbar}\big) \\
= \; & \hat{T}\big(\hat{A}(G \otimes \mathrm{e}_{\otimes}^{F}) \otimes \mathrm{e}_{\otimes}^{iF/\hbar}\big) \\
= \; & \hat{s}_0 \hat{T}\big(G \otimes \mathrm{e}_{\otimes}^{iF/\hbar}\big) - \hat{T}\big((S_0 + F, G) \otimes \mathrm{e}_{\otimes}^{iF/\hbar}\big) \\
- \; & \frac{i}{2\hbar} \hat{T}\big((S_0 + F, S_0 + F) \otimes G \otimes \mathrm{e}_{\otimes}^{iF/\hbar}\big) \\
= \; & \hat{s}_0 T\big([G + D(G \otimes \mathrm{e}_{\otimes}^{F})] \otimes \mathrm{e}_{\otimes}^{iF/\hbar}\big) - T\big((S_0 + F, G) \otimes \mathrm{e}_{\otimes}^{iF/\hbar}\big) \\
- \; & \frac{i}{2\hbar} T\big((S_0 + F, S_0 + F) \otimes [G + D(G \otimes \mathrm{e}_{\otimes}^{F})] \otimes \mathrm{e}_{\otimes}^{iF/\hbar}\big) \\
= \; & T\big(A(G \otimes \mathrm{e}_{\otimes}^{F}) \otimes \mathrm{e}_{\otimes}^{iF/\hbar}\big) + T\big(A[D(G \otimes \mathrm{e}_{\otimes}^{F}) \otimes \mathrm{e}_{\otimes}^{F}] \otimes \mathrm{e}_{\otimes}^{iF/\hbar}\big) \\
+ \; & T\big((S_0 + F, D(G \otimes \mathrm{e}_{\otimes}^{F})) \otimes \mathrm{e}_{\otimes}^{iF/\hbar}\big) . \qquad (405)
\end{aligned}$$

It follows that the anomalies for the old and new prescription are related by

$$\begin{aligned}
& A\big(G \otimes \mathrm{e}_{\otimes}^{F}\big) + A\big(D(G \otimes \mathrm{e}_{\otimes}^{F}) \otimes \mathrm{e}_{\otimes}^{F}\big) + \big(S_0 + F, D(G \otimes \mathrm{e}_{\otimes}^{F})\big) \\
= \; & D\big(\hat{A}(G \otimes \mathrm{e}_{\otimes}^{F}) \otimes \mathrm{e}_{\otimes}^{F}\big) + \hat{A}\big(G \otimes \mathrm{e}_{\otimes}^{F}\big) . \qquad (406)
\end{aligned}$$

The trick is again to evaluate this at order $\hbar^m$, and use that $D(G \otimes \mathrm{e}_{\otimes}^{F})$ and and $A(G \otimes \mathrm{e}_{\otimes}^{F})$ themselves are of order $\hbar^m$. This gives the equality

$$A^m\big(G \otimes \mathrm{e}_{\otimes}^{F}\big) + \big(S_0 + F, D^m(G \otimes \mathrm{e}_{\otimes}^{F})\big) = \hat{A}^m\big(G \otimes \mathrm{e}_{\otimes}^{F}\big) . \qquad (407)$$

Now, if $G = \int \gamma \wedge \mathbf{J}$, and if $F = \lambda S_1 \otimes \lambda^2 S_2$, then it follows from the above equation that

$$A^m\big(G \otimes \mathrm{e}_{\otimes}^{\lambda S_1 + \lambda^2 S_2}\big) + \hat{s} D^m\big(G \otimes \mathrm{e}_{\otimes}^{\lambda S_1 + \lambda^2 S_2}\big) = \hat{A}^m\big(G \otimes \mathrm{e}_{\otimes}^{\lambda S_1 + \lambda^2 S_2}\big) . \qquad (408)$$

Furthermore, it follows from the definition of D that $D(G \otimes \mathrm{e}_{\otimes}^{\lambda S_1 + \lambda^2 S_2})$ is equal to $-\int \gamma \wedge j^m$. Furthermore, if γ is a closed 1-form, we have shown above that $\hat{s} \int \gamma \wedge j^m = \int \gamma \wedge h^m$. By eq. (400) and our definition of D, we therefore have

$$\hat{s} D^m\big(G \otimes \mathrm{e}_{\otimes}^{\lambda S_1 + \lambda^2 S_2}\big) = -A^m\big(G \otimes \mathrm{e}_{\otimes}^{\lambda S_1 + \lambda^2 S_2}\big) . \qquad (409)$$

Consequently, we have shown that

$$\hat{A}^m\big(G \otimes \mathrm{e}_{\otimes}^{\lambda S_1 + \lambda^2 S_2}\big) = 0 . \qquad (410)$$

Therefore, our redefinition of the time ordered products has already removed the anomaly $\hat{A}(G \otimes \mathrm{e}_{\otimes}^{F})$ in the case when γ is a closed 1-form, and f is a constant. We now drop the carret from our notation for the newly defined time-ordered products and the corresponding anomaly.

We may then assume that eq. (410) holds for A^m. For the quantities defined in eq. (411), this means that

$$0 = \int_{M^n} \mathcal{A}^m_n(x_1,\dots,x_n)\gamma(x_1)\,dx_1\dots dx_n\,. \tag{411}$$

for any closed 1-form γ, and any n. Lemma (9) now implies that we we may write $\mathcal{A}^m_n$ as

$$\mathcal{A}^m_n(x_1,\dots,x_n) = d_1\mathcal{B}^m_{n/1}(x_1,\dots,x_n) + \sum_{k=2}^{n} d_k\,\mathcal{B}^m_{n/k}(x_1,\dots,x_n) \tag{412}$$

Here, the $\mathcal{B}_{m,n/k}$ are now a local covariant functional of $(\Phi,\Phi^\ddagger)$ in the space $\mathbf{P}^{2/4/\dots4}(M^n)$ for $k=1$, and in the space $\mathbf{P}^{3/4/\dots3/\dots4}(M^n)$ for $k\geq 2$. By elementary manipulations using δ-functions, using that the $\mathcal{B}^m_{n/k}$ are supported on the diagonal and have dimension +2 and ghost number +2, we may shift the derivatives between the terms in the sum if necessary and thereby achieve that

$$\mathcal{B}^m_{n/1}(x_1,\dots,x_n) = \text{const.}\,\mathbf{K}_1(x_1)\,\delta(x_1,\dots,x_n)\,, \tag{413}$$

up to an irrelevant total d_1-derivative. Next, we define for products with n arguments containing 1 factor of $\mathbf{K}_1\in\mathbf{P}^2_2$ [see eq. (76)] and $n-1$ factors of $\mathbf{L}_1\in\mathbf{P}^4_0$ by

$$D_n(\mathbf{K}_1(x_1)\otimes\mathbf{L}_1(x_2)\cdots\otimes\mathbf{L}_1(x_n)) := -\hbar^m\mathcal{B}^m_{n/1}(x_1,\dots,x_n)\,. \tag{414}$$

We redefine the time-ordered products with $n+1$ factors, containing 1 factor of $\mathbf{J}_1\in\mathbf{P}^3_1$, one factor of $\mathcal{O}_1\in\mathbf{P}^3_1$ [see eq. (391)], and $n-2$ factors of $\mathbf{L}_1\in\mathbf{P}^4_0$ by

$$D_n(\mathbf{J}_1(x_1)\otimes\mathbf{L}_1(x_2)\cdots\otimes\mathcal{O}_1(x_k)\cdots\otimes\mathbf{L}_1(x_n)) := i\hbar^{m+1}\mathcal{B}^m_{n/k}(x_1,\dots,x_n)\,. \tag{415}$$

By going through the same steps as above in T12a, we find that the new anomaly $\hat{A}(G\otimes \mathrm{e}^F_\otimes)$ after the above redefinition effected by these D's is now

$$\hat{A}\big(G\otimes\mathrm{e}^F_\otimes\big) = A\big(G\otimes\mathrm{e}^F_\otimes\big) - D\big((S_0+F,G)\otimes\mathrm{e}^F_\otimes\big) + \frac{i}{2\hbar}D\big((S_0+F,S_0+F)\otimes G\otimes\mathrm{e}^F_\otimes\big)\,. \tag{416}$$

Now, it can be seen that, because of the first redefinition (414),

$$D\big((S_0+F,G)\otimes\mathrm{e}^F_\otimes\big) = \hbar^m\sum_{n\geq 0}\frac{\lambda^n}{n!}\int d_1\,\mathcal{B}^m_{n/1}(x_1,\dots,x_n)\,\gamma(x_1)\,f(x_1)\dots f(x_n)\,dx_1\dots dx_n\,, \tag{417}$$

using $(S_0,\mathbf{J}_1) = d\mathbf{K}_1+\dots$. It follows from the second redefinition (415) that

$$\begin{aligned}&\frac{i}{2\hbar}D\big(G\otimes(S_0+F,S_0+F)\otimes\mathrm{e}^F_\otimes\big) \qquad (418)\\ = \;& -\hbar^m\sum_{n\geq 0}\sum_{k=2}^{n}\frac{\lambda^n}{n!}\int d_k\,\mathcal{B}^m_{n/k}(x_1,\dots,x_n)\,\gamma(x_1)\,f(x_1)\dots f(x_k)\dots f(x_n)\,dx_1\dots dx_n\,,\end{aligned}$$

using $(S_0, \mathbf{L}_1) = d\mathcal{O}_1$. Thus, taking the $O(\hbar^m)$-part of eq. (416), using eq. (413), we find that the new anomaly $\hat{A}^m(G \otimes \mathrm{e}_\otimes^F) = 0$. Thus, the anomaly for the new time-ordered products vanishes at order $\hbar^m$ and to all orders in λ. We continue this process by redefining the time ordered products to the next order in $\hbar$, and remove the anomaly $A^{m+1}(G \otimes \mathrm{e}_\otimes^F)$. Since we can do this for all m, we see that we can satisfy T12b above by a suitable redefinition of the time-ordered products. $\square$

4.4.3 Proof of T12c

Let $\Psi = \prod \Theta_{s_i}(F, \mathcal{D}F, \mathcal{D}^2 F, \dots)$ be the gauge-invariant expression of form-degree p under consideration, where Θ_s are invariant polynomials of the Lie-algebra, so that in particular $\hat{s}\Psi = 0$. Let α be a $(4-p)$-form, and let $G = \int \alpha \wedge \Psi$. The satisfaction of the Ward identity T12c means that anomaly in eq. (395) can be removed, where G in that equation is now $\int \alpha \wedge \Psi$. As in the proofs of T12a, T12b, one first proves the consistency condition

$$\hat{s}A^m\big(G \otimes \mathrm{e}_\otimes^{\lambda S_1 + \lambda^2 S_2}\big) = 0\,, \tag{419}$$

where m is the first order in $\hbar$ where the anomaly occurs, and where α is now arbitrary. This condition is again of cohomological nature. As in T12b, it may be used to show that the anomaly can be removed, at n-th order in λ, by a redefinition of the time ordered products with 1 factor of Ψ_0 and n factors of $\mathbf{L}_1$, and by the time ordered products with 1 factor of Ψ_0, 1 factor of $\mathcal{O}_1$ [see eq. (391)] and $n-1$ factors of $\mathbf{L}_1$. The details of these arguments are completely analogous to those given above in the proofs of T12a and T12b, so we omit them here. $\square$

4.5 Formal BRST-invariance of the S-matrix

We consider the adiabatically switched S-matrix $\mathcal{S}(F) = T(\mathrm{e}_\otimes^{iF/\hbar})$ associated with the cut-off interaction $F = \int_M \{\lambda f \mathbf{L}_1 + \lambda^2 f^2 \mathbf{L}_2\}$, where f is a smooth switching function of compact support. Let Q_0 be the free BRST-charge operator. It follows from the definition $\mathcal{S}(F) = T(\mathrm{e}_\otimes^{iF/\hbar})$ and the Ward-identities T12a [see eq. (322)] that

$$[Q_0, \mathcal{S}(F)] = -\frac{1}{2} T\Big((S_0 + F, S_0 + F) \otimes \mathrm{e}_\otimes^{iF/\hbar}\Big) \quad \mathrm{mod}\ \mathcal{J}_0\,. \tag{420}$$

Now consider a sequence of cutoff functions such that $f \to 1$ sufficiently rapidly, i.e., the "adiabatic limit". Then it follows that $S_0 + F \to S$, and consequently that $(S_0 + F, S_0 + F) \to (S, S) = 0$. Thus, formally, $T((S_0 + F, S_0 + F) \otimes \mathrm{e}_\otimes^{iF/\hbar}) \to 0$. Furthermore, formally, $\mathcal{S}(F)$ converges to the true S-matrix $\mathcal{S}$. Consequently, assuming that all these limits exist, we would have

$$[Q_0, \mathcal{S}] = 0 \quad \mathrm{mod}\ \mathcal{J}_0\ (\mathrm{FORMALLY})\,. \tag{421}$$

As we have already said, the adiabatic limit does not appear to exist for pure Yang-Mills theory in Minkowski spacetime, and there is even less reason to believe that it ought to exist in generic

curved spacetimes. Therefore, the above statement concerning the BRST-invariance of the S-matrix is most likely only a formal statement, unlike the other results in this paper. We have nevertheless mentioned it, because such a condition is often taken to be as the definition of gauge-invariance at the perturbative level in less rigorous treatments of quantum gauge field theories in flat spacetime.

4.6 Proof that $d\mathbf{J}_I = 0$

As above, consider the cutoff interaction $F = \int_M \{\lambda f \mathbf{L}_1 + \lambda^2 f^2 \mathbf{L}_2\}$, where f is a smooth switching function of compact support, which is equal to one on some time-slice $M_T = (-T, T) \times \Sigma$. The desired identity $d\mathbf{J}(x)_I$ will follow if we can show that, in the sense of formal power series,

$$0 = d\mathbf{J}(x)_F = \sum_n \frac{i^n}{\hbar^n n!} R_n(d\mathbf{J}(x); F^{\otimes n}), \quad x \in M_T \tag{422}$$

modulo $\mathcal{J}_0$ for any such cutoff function f. Expanding the retarded products in terms of time ordered products gives the equivalent relation

$$T\left(d\mathbf{J}(x) \otimes \mathrm{e}_\otimes^{iF/\hbar}\right) = 0 \quad \mathrm{mod}\ \mathcal{J}_0 \text{ forall } x \in M_T, \tag{423}$$

which is again to be understood in the sense of formal power series. At the level of classical fields, we have

$$d\mathbf{J}(x) = (S_0 + F, \Phi(x)) \cdot (\Phi^\ddagger(x), S_0 + F) \quad \text{forall } x \in M_T. \tag{424}$$

Hence, (423) is equivalent to the equation

$$\begin{aligned} T\left(d\mathbf{J}_0(x) \otimes \mathrm{e}_\otimes^{iF/\hbar}\right) \ &= \ -T\left(\left\{\hat{s}_0\Phi(x) \cdot (\Phi^\ddagger(x), F) + \hat{s}_0\Phi^\ddagger(x) \cdot (\Phi(x), F)\right\} \otimes \mathrm{e}_\otimes^{iF/\hbar}\right) \\ &\quad -T\left(\left\{(F, \Phi(x)) \cdot (\Phi^\ddagger(x), F)\right\} \otimes \mathrm{e}_\otimes^{iF/\hbar}\right) \quad \mathrm{mod}\ \mathcal{J}_0. \end{aligned} \tag{425}$$

We claim that this equation can be satisfied as a consequence of our Ward identity T12a by a redefinition of the time-ordered products. In fact, we shall now show that our Ward identity

T12a can even be used to prove the following stronger identity:

$$
\begin{aligned}
&\sum_{I_1\cup\cdots\cup I_t=\underline{n}} \left(\frac{i}{\hbar}\right)^t T_{t+1}(d\mathbf{J}_0(y)\otimes \mathbf{L}_{|I_1|}(X_{I_1})\otimes\dots\mathbf{L}_{|I_t|}(X_{I_t})) = \\
&\qquad -\sum_{I_1\cup\cdots\cup I_t=\underline{n}} \left(\frac{i}{\hbar}\right)^{t-1} \sum_{i=1}^{t} T_t\Big(\mathbf{L}_{|I_1|}(X_{I_1})\otimes\dots \\
&\qquad \otimes\Big\{\hat{s}_0\Phi(y)\cdot(\Phi^\ddagger(y),\mathbf{L}_{|I_i|}(X_{I_i})) + \hat{s}_0\Phi^\ddagger(y)\cdot(\Phi(y),\mathbf{L}_{|I_i|}(X_{I_i}))\Big\}\otimes\dots\mathbf{L}_{|I_t|}(X_{I_t})\Big) \\
&\qquad -\sum_{I_1\cup\cdots\cup I_t=\underline{n}} \left(\frac{i}{\hbar}\right)^{t-2} \sum_{1\le i<j\le t} T_{t-1}\Big(\mathbf{L}_{|I_1|}(X_{I_1})\otimes\dots \\
&\qquad (\mathbf{L}_{|I_i|}(X_{I_i}),\Phi^\ddagger(y))\cdot(\Phi(y),\mathbf{L}_{|I_j|}(X_{I_j}))\otimes\dots\mathbf{L}_{|I_t|}(X_{I_t})\Big) \qquad (426)
\end{aligned}
$$

modulo $\mathcal{J}_0$. This identity implies (423) as may be seen by multiplying each term by $\lambda^n/n!$, integrating against $f(x_1),\dots,f(x_n)$, and summing over n. Thus, it remains to be seen that (426) follows from the Ward identity T12a. For $n=0$, we get the condition $T_1(d\mathbf{J}_0(y))=0$, which is just the condition of current conservation in the free theory and hence is satisfied. For $n>0$, we proceed inductively. This shows that, at the order considered, the failure of (426) to be satisfied is of the form $T_1(\alpha_n(y,x_1,\dots,x_n))$, where $\alpha_n(y,x_1,\dots,x_n)$ is a local covariant functional that is supported on the total diagonal. We now show that we can set this quantity to 0. To do this, we pick a testfunction $h\in C^\infty(M)$ with the following properties: $h(y)=1$ in an open neighborhood of $\{x_1,\dots,x_n\}$, $h(y)=0$ towards the future of Σ_+, and towards the past of Σ_-, where $\Sigma_\pm$ are Cauchy surfaces in the future/past of $\{x_1,\dots,x_n\}$. We may thus write $dh=\gamma_+-\gamma_-$, where $\gamma_\pm$ are 1-forms that are supported in the future/past of $\{x_1,\dots,x_n\}$. Now, from $Q_0=\int_M T_1(\mathbf{J}_0)\wedge\gamma_\pm$, and from the causal factorization of the time-ordered products, we have

$$
\begin{aligned}
&\int_M h(y)T_{t+1}(d\mathbf{J}_0(y)\otimes\mathbf{L}_{|I_1|}(X_{I_1})\otimes\dots\mathbf{L}_{|I_t|}(X_{I_t}))\,dy \\
=\;& [Q_0,T_t(\mathbf{L}_{|I_1|}(X_{I_1})\otimes\dots\mathbf{L}_{|I_t|}(X_{I_t}))] = i\hbar\hat{s}_0 T_t(\mathbf{L}_{|I_1|}(X_{I_1})\otimes\dots\mathbf{L}_{|I_t|}(X_{I_t}))\,, \qquad (427)
\end{aligned}
$$

where the last equation is modulo $\mathcal{J}_0$. We also have

$$
\int_M h(y)(\mathcal{O}(x_i),\Phi^\ddagger(y))\cdot(\Phi(y),\mathcal{O}(x_j))\,dy = (\mathcal{O}(x_i),\mathcal{O}(x_j)) \qquad (428)
$$

for any $\mathcal{O}$. It follows from these equations that if we integrate (426) against $h(y)$, then we get an identity follows from the known Ward identity T12a. Stated differently, because $h(y)=1$ in a neighborhood of $\{x_1,\dots,x_n\}$, and because the failure α_n of (426) to hold is supported on the total diagonal, it must satisfy

$$
\int_M \alpha_n(y,x_1,\dots,x_n)\,dy = 0 \mod \mathcal{J}_0. \qquad (429)
$$

By lemma 9, it hence follows that there exists a local covariant β_n supported on the total diagonal such that $d_y \beta_n(y, x_1, \dots, x_n) = \alpha_n(y, x_1, \dots, x_n)$, where β_n is a 3-form in the y-entry, and a 4-form in each x_i-entry, and where d_y is the exterior differential acting on the y-variable. We may now redefine time ordered products with one factor of $\mathbf{J}_0(y)$ and n factors of $\mathbf{L}_1(x_i), i = 1, \dots, n$ by taking $D_{n+1}(\mathbf{J}_0(y) \otimes \mathbf{L}_1(x_1) \otimes \cdots \otimes \mathbf{L}_1(x_n)) := \beta_n(y, x_1, \dots, x_n)$. Then the redefined time-ordered products satisfy (426).

4.7 Proof that $Q_I^2 = 0$

We know from the previous subsection that the interacting BRST-current is conserved, $d\mathbf{J}(x)_I = 0$ for any x, or equivalently, $d\mathbf{J}(x)_F = 0$ for any x in a domain $M_T = (-T, T) \times \Sigma$ where the function f in $F = \int \{\lambda f \mathbf{L}_1 + \lambda^2 f^2 \mathbf{L}_2\}$ is equal to 1. Thus, the definition of the interacting BRST-charge $Q_I = \int \gamma \wedge \mathbf{J}_I$ is independent of the choice of the compactly supported closed 1-form γ dual to the Cauchy surface Σ. Using the Bogoliubov formula for the interacting field operators, the desired equality $Q_I^2 = 0$ is equivalent to the equation

$$
\begin{aligned}
&0 = Q_F^2 = \left(\int \gamma(x) \wedge \mathbf{J}(x)_F \right)^2 = \\
&\frac{1}{2} \sum_{n,m} \frac{i^{n+m}}{\hbar^{n+m} n! m!} \int \left[R_n(\mathbf{J}(x); F^{\otimes n}), R_m(\mathbf{J}(y); F^{\otimes m}) \right] \gamma(x) \gamma(y) \, dx dy
\end{aligned}
\tag{430}
$$

modulo $\mathcal{J}_0$, where γ is now chosen to be supported in M_T. Note that, as usual, we mean the graded commutator, which is actually the anti-commutator in the above expression. Now, because the interacting BRST-charge Q_F as defined using the cutoff interaction F is independent upon the choice of the compactly supported closed 1-form in γ dual to Σ, we may write the interacting BRST-charge either as $Q_F = \int \gamma^{(1)} \wedge \mathbf{J}_F$, or as $Q_F = \int \gamma^{(2)} \wedge \mathbf{J}_F$. We may therefore alternatively write

$$
Q_F^2 = \frac{1}{2} \sum_n \frac{i^n}{\hbar^n n!} \int R_{n+1} \Big(\mathbf{J}(x); \mathbf{J}(y) \otimes F^{\otimes n} \Big) \gamma^{(1)}(x) \gamma^{(2)}(y) \, dx dy + (1 \leftrightarrow 2) \,, \tag{431}
$$

where we have also used the GLZ-formula (266). We now make a particular choice for $\gamma^{(1)}$ and $\gamma^{(2)}$ that will facilitate the evaluation of this expression. We choose $\gamma^{(1)} = dh^{(1)} + dh^{(2)}$, where $h^{(1)}$ and $h^{(2)}$ are smooth scalar functions with the following properties: (a) the support of $h^{(1)}$ is compact, (b) $h^{(1)} = 1$ on the support of $\gamma^{(2)}$, (c) the support of $h^{(2)}$ is contained in the causal past of the support of $\gamma^{(2)}$. Due to these support properties and the causal support properties of the retarded products, the above expression can then be written as

$$
Q_F^2 = -\frac{1}{2} \sum_n \frac{i^n}{\hbar^n n!} R_{n+1} \Big(d\mathbf{J}(x); \mathbf{J}(y) \otimes F^{\otimes n} \Big) \, h^{(1)}(x) \gamma^{(2)}(y) \, dx dy \tag{432}
$$

Below, we will show that, for any $x, y \in M_T$, the following identity is a consequence of the Ward-identity T12b:

$$
\begin{aligned}
&R\Big(d\mathbf{J}(x);\mathbf{J}(y)\otimes \mathrm{e}_\otimes^{iF/\hbar}\Big) = \\
&i\hbar R\Big(\Big\{(S_0+F,\Phi(x))\cdot(\Phi^\ddagger(x),\mathbf{J}(y))+(S_0+F,\Phi^\ddagger(x))\cdot(\Phi(x),\mathbf{J}(y))\Big\};\mathrm{e}_\otimes^{iF/\hbar}\Big) \quad \mathrm{mod}\, \mathcal{J}_0\,. \qquad (433)
\end{aligned}
$$

We now apply this identity and use that $h^{(1)} = 1$ on the support of $\gamma^{(2)}$. Then we obtain

$$
Q_F^2 = \frac{i\hbar}{2}\int R\Big((S,\mathbf{J}(x));\mathrm{e}_\otimes^{iF/\hbar}\Big)\,\gamma^{(2)}(x)\,dx\,, \tag{434}
$$

again, modulo $\mathcal{J}_0$. However, $\hat{s}\mathbf{J} = d\mathbf{K}$, so using T11, the right side vanishes by $d\gamma^{(2)} = 0$. Thus, we have proved $Q_F^2 = 0$ modulo $\mathcal{J}_0$, and it remains to prove eq. (433). That equation can be written equivalently in terms of time ordered products

$$
\begin{aligned}
&T\Big(d\mathbf{J}(x)\otimes\mathbf{J}(y)\otimes \mathrm{e}_\otimes^{iF/\hbar}\Big) \qquad (435)\\
&= i\hbar T\Big(\Big\{(S_0+F,\Phi(x))\cdot(\Phi^\ddagger(x),\mathbf{J}(y))+(S_0+F,\Phi^\ddagger(x))\cdot(\Phi(x),\mathbf{J}(y))\Big\}\otimes\mathrm{e}_\otimes^{iF/\hbar}\Big) \quad \mathrm{mod}\, \mathcal{J}_0\,,
\end{aligned}
$$

using the formulae relating time-ordered and retarded products given above. We will prove it in this form. Using eq. (73), the eq. (435) may be written alternatively as

$$
\begin{aligned}
&T\Big(d\mathbf{J}_0(x)\otimes\mathbf{J}(y)\otimes \mathrm{e}_\otimes^{iF/\hbar}\Big) = \qquad (436)\\
&-T\Big(\Big\{\hat{s}_0\Phi(x)\cdot(\Phi^\ddagger(x),F)+(\Phi\leftrightarrow\Phi^\ddagger)\Big\}\otimes\mathbf{J}(y)\otimes\mathrm{e}_\otimes^{iF/\hbar}\Big)\\
&-T\Big(\Big\{(F,\Phi(x))\cdot(\Phi^\ddagger(x),F)+(\Phi\leftrightarrow\Phi^\ddagger)\Big\}\otimes\mathbf{J}(y)\otimes\mathrm{e}_\otimes^{iF/\hbar}\Big)\\
&+i\hbar T\Big(\Big\{\hat{s}_0\Phi(x)\cdot(\Phi^\ddagger(x),\mathbf{J}(y))+(F,\Phi(x))\cdot(\Phi^\ddagger(x),\mathbf{J}(y))+(\Phi\leftrightarrow\Phi^\ddagger)\Big\}\otimes\mathrm{e}_\otimes^{iF/\hbar}\Big) \quad \mathrm{mod}\, \mathcal{J}_0.
\end{aligned}
$$

We will now show that this equation can be satisfied as a consequence of our Ward-identity T12b. To prove this identity, we employ the same technique as in the previous subsection. We first formulate a set of stronger identities that will imply . This set of conditions is completely analogous to eqs. (426), with the difference that in eq. (426), we replace $\mathbf{L}_i(X)$ everywhere by $\mathbf{L}_i(X)+\tau\mathbf{J}_i(y,X)$, and expand the resulting set of equations to first order in τ. As in the proof of eqs. (426), the resulting equations are established inductively in n. For $n = 0$ the identity can be verified directly using the definitions made in free gauge theory. Inductively, the resulting equations will then be violated at order n by a potential "anomaly" term of the form $T_1(\alpha_n(x,y,x_1,\dots,x_n))$, where α_n is now an element of $\mathbf{P}^{4/3/4/\dots/4}(M^{n+2})$. As in the treatment of eq. (426), the Ward identity T12b then implies that

$$
\int_M \alpha_n(x,y,x_1,\dots,x_n)\,dx = 0 \tag{437}
$$

while the GLZ-identity, together with the fact that $d\mathbf{J}_I = 0$ can be seen to imply the relation

$$\int_M d_y \alpha_n(x,y,x_1,\dots,x_n)\, dx_1 \dots dx_n = 0\,. \tag{438}$$

Eqs.(437) and (438) can now be used to show that the time-ordered products can be redefined, if necessary, to remove the anomaly α_n. By the same argument as in the previous subsection, the first identity (437) implies that

$$\alpha_n(x,y,x_1,\dots,x_n) = d_x \delta_n(x,y,x_1,\dots,x_n) \tag{439}$$

for some $\delta_n \in \mathbf{P}^{3/3/4/\dots/4}(M^{n+2})$. We would like to redefine the time-ordered products using the quantity D_n (see sec. 3.6)

$$D_{n+2}(\mathbf{J}_0(x) \otimes \mathbf{J}_0(y) \otimes \mathbf{L}_1(x_1) \cdots \otimes \mathbf{L}_1(x_n) := \delta_n(x,y,x_1,\dots,x_n)\,. \tag{440}$$

In view of eq. (439), this would remove the anomaly. However, it is not clear that we can make this redefinition, because the time-ordered products with two free BRST-currents at x and y must be anti-symmetric in x and y, and this need not be the case for δ_n in (439). We will circumvent this problem by using a modified $\hat{\delta}_n$ in eq. (440) to redefine the time-ordered products with 2 currents. To construct the modified $\hat{\delta}_n$, we consider the quantity

$$\beta(\gamma^{(1)},\gamma^{(2)}) = \int \delta_n(x,y,z_1,\dots,z_n)\gamma^{(1)}(x)\gamma^{(2)}(y)\, dxdydz_1 \dots dz_n + (1 \leftrightarrow 2)\,, \tag{441}$$

where $\gamma^{(1)},\gamma^{(2)}$ are now arbitrary 1-forms of compact support. β is evidently closely related to the symmetric part of δ_n, which we would like to be zero. From eq. (438), we have $\beta(dh^{(1)},dh^{(2)}) = 0$ for any pair of compactly supported scalar functions $h^{(1)},h^{(2)}$. As we shall show presently, this implies that we can write

$$\beta(\gamma^{(1)},\gamma^{(2)}) = C(d\gamma^{(1)},\gamma^{(2)}) + (1 \leftrightarrow 2) \tag{442}$$

where C has a distributional kernel $C \in \mathbf{P}^{2/3}(M^2)$. We now define

$$\hat{\delta}_n(x,y,z_1,\dots,z_n) = \delta_n(x,y,z_1,\dots,z_n) - d_x C(x,y)\delta(y,z_1,\dots,z_n) - (x \leftrightarrow y)\,, \tag{443}$$

which is manifestly anti-symmetric in x,y. We use this new $\hat{D}_n$ in order to redefine the time-ordered products with 2 currents as in eq. (440) instead of the old D_n. Evidently, the new time ordered product is now anti-symmetric in x,y. Furthermore, as a consequence of eq. (442), the new anomaly for the redefined time-ordered products $\hat{\alpha}_n$ satisfies

$$\int \hat{\alpha}_n(x,y,z_1,\dots,z_n)\, dz_1 \dots dz_n = 0\,. \tag{444}$$

It follows from this equation that

$$\hat{\alpha}_n(x,y,z_1,\dots,z_n) = \sum_{l=1}^{n} d_l \delta_{n/l}(x,y,z_1,\dots,z_n) \quad d_l = dz_l \wedge \frac{\partial}{\partial z_l} \tag{445}$$

for some $\delta_{n/l} \in \mathbf{P}^{4/3/4/\dots/3\dots/4}(M^{n+2})$. We use these quantities to make a final redefinition of the time-ordered products. We have

$$\hat{s}_0\Phi(x_1)\cdot(\Phi^\ddagger(x_1),\mathbf{L}_1(x_2))+\hat{s}_0\Phi^\ddagger(x_1)\cdot(\Phi(x_1),\mathbf{L}_1(x_2)) = d_1\mathbf{J}_1(x_1)\delta(x_1,x_2)+d_2\Sigma_1(x_1,x_2) \tag{446}$$

for some $\Sigma_1 \in \mathbf{P}^{3/3}(M^2)$. We redefine the time-ordered products involving these quantities using the quantities (see sec. 3.6)

$$D_{n+1}(\mathbf{J}_0(x)\otimes\mathbf{L}_1(z_1)\cdots\otimes\Sigma_1(y,z_l)\otimes\dots\mathbf{L}_1(z_n)) := \delta_{n/l}(x,y,z_1,\dots,z_n))\,. \tag{447}$$

This final redefinition then removes the anomaly $\hat{\alpha}_n$. $\square$

It remains to prove eq. (442). We formulate this result as a lemma:

Lemma 10. Let $\beta \in \mathbf{P}^{3/3}(M^2)$ such that $\beta(dh^{(1)},dh^{(2)}) = 0$ for any pair of compactly supported scalar functions $h^{(1)},h^{(2)}$. Then β can be written in the form (442) for some $C \in \mathbf{P}^{2/3}(M^2)$.

Proof: β is of the form

$$\beta(\gamma^{(1)},\gamma^{(2)}) = \int_M dx \sum_{m=0}^{p} \beta^{\mu\nu_1\dots\nu_m\sigma}\gamma^{(1)}_\mu \nabla_{\nu_1}\cdots\nabla_{\nu_m}\gamma^{(2)}_\sigma\,, \tag{448}$$

where β are tensor fields that are locally constructed out of g,∇, and $\Phi,\Phi^\ddagger$. We claim that the condition $\beta(dh^{(1)},dh^{(2)}) = 0$ and the symmetry of β implies that β can be put into the form (442). Since the commutator of two derivatives gives a Riemann tensor, we may assume that each tensor β in the sum in (448) is symmetric under the exchange of the indices $\nu_1,\dots,\nu_m$,

$$\beta^{\mu\nu_1\dots\nu_m\sigma} = \beta^{\mu(\nu_1\dots\nu_m)\sigma}\,. \tag{449}$$

Now consider the contribution to (448) with the highest number of derivatives, $m = p$. By varying $\beta(dh^{(1)},dh^{(2)}) = 0$ with respect to $h^{(1)},h^{(2)}$ there follows the additional symmetry

$$\beta^{(\mu\nu_1\dots\nu_p\sigma)} = 0\,. \tag{450}$$

Consider now the vector field defined by

$$B^\mu = \beta^{\mu\nu_1\dots\nu_p\sigma}\nabla_{\nu_1}\cdots\nabla_{\nu_p}\gamma_\sigma\,. \tag{451}$$

Using the symmetry property (449), this may be rewritten as

$$\begin{aligned} B^\mu &= \beta^{\mu\nu_1\dots\nu_p\sigma}\nabla_{\nu_1}\cdots\nabla_{[\nu_p}\gamma_{\sigma]} \\ &\quad +\beta^{\mu(\nu_1\dots\nu_p\sigma)}\nabla_{\nu_1}\cdots\nabla_{\nu_p}\gamma_\sigma\,. \end{aligned} \tag{452}$$

Then, using the symmetry (450), this may further be written as

$$
\begin{aligned}
B^{\mu} &= \beta^{\mu\nu_1\ldots\nu_p\sigma}\nabla_{\nu_1}\cdots\nabla_{[\nu_p}\gamma_{\sigma]} \\
&\quad -\frac{2}{p+2}\beta^{\sigma(\mu\nu_1\ldots\nu_p)}\nabla_{\nu_1}\cdots\nabla_{[\nu_p}\gamma_{\sigma]} \\
&\quad -\frac{2(p+1)}{p+2}\nabla_{\nu}\left\{\beta^{\mu(\nu\alpha_1\ldots\alpha_{p-1}\sigma)}\nabla_{\alpha_1}\cdots\nabla_{\alpha_{p-1}}\gamma_{\sigma}-(\mu\leftrightarrow\nu)\right\} \qquad (453)\\
&\quad +\text{terms with } (p-1) \text{ derivatives on } \gamma_{\sigma}\,. \qquad (454)
\end{aligned}
$$

Now put $\gamma=\gamma^{(2)}$ in this equation, contract both sides with $\gamma^{(1)}$, and integrate, to obtain an expression for the highest derivative term in β. Using this expression, we find that $\beta(\gamma^{(1)},\gamma^{(2)})$ is given by a sum of terms each of which contains either $\nabla_{[\mu}\gamma^{(1)}_{\nu]}$ or $\nabla_{[\mu}\gamma^{(2)}_{\nu]}$, or which contains at most derivative terms of order $p-1$. Consequently, using the symmetry of β, we can write

$$\beta(\gamma^{(1)},\gamma^{(2)})=C(d\gamma^{(1)},\gamma^{(2)})+C(d\gamma^{(2)},\gamma^{(1)})+R_{p-1}(\gamma^{(1)},\gamma^{(2)}), \qquad (455)$$

where R_{p-1} stands for a remainder term of the form (448) containing at most $p-1$ derivatives, and where C is also of the form (448). If we now take $\gamma^{(1)}=dh^{(1)}$, and $\gamma^{(2)}=dh^{(2)}$ in eq. (455), and use $\beta(dh^{(1)},dh^{(2)})=0$, then we see that R_{p-1} again satisfies $R_{p-1}(dh^{(1)},dh^{(2)})=0$. Thus, we may repeat the arguments just given for R_{p-1} and conclude that β can be written as in eq. (455) with a new C, and a remainder R_{p-2} containing at most $p-2$ derivatives. Thus, further repeating this procedure, we find that (455) must hold for some C and a remainder of the form $R_0(\gamma^{(1)},\gamma^{(2)})=\int\varepsilon\gamma^{(1)}_{\mu}r^{\mu\nu}\gamma^{(2)}_{\nu}$.

Now, R_0 is symmetric, so $r^{[\mu\nu]}=0$. Furthermore, we have $R_0(dh^{(1)},dh^{(2)})=0$ for all compactly supported $h^{(1)},h^{(2)}$. Varying this equation with respect to $h^{(2)}$, we get $0=\nabla^{\mu}(r_{\mu\nu}\nabla^{\nu}h^{(1)})$. Now, pick a point $x\in M$, and choose $h^{(1)}$ so that $h^{(1)}(x)=0$. Then it follows that $r_{\mu\nu}\nabla^{\mu}\nabla^{\nu}h^{(1)}=0$ at x. Because $\nabla^{\mu}\nabla^{\nu}h^{(1)}$ is an arbitrary symmetric tensor at x, it follows that $r^{(\mu\nu)}=0$, and therefore that $r^{\mu\nu}=0$, thus proving the desired decomposition (442). This completes the proof. □

4.8 Proof that $[Q_I,\Psi_I]=0$ when Ψ is gauge invariant

Here we show that the Ward identity T12c implies $[Q_I,\Psi_I(x)]=0$ modulo $\mathcal{J}_0$, whenever $\Psi\in\mathbf{P}(M)$ is a strictly gauge invariant operator of ghost number 0, i.e., $\Psi=\prod\Theta_{s_i}(F,\mathcal{D}F,\ldots,\mathcal{D}^{k_i}F)$. As in the proof given in the previous subsection, this property will follow from the identity

$$
\begin{aligned}
&T\left(d\mathbf{J}_0(x)\otimes\Psi(y)\otimes e_{\otimes}^{iF/\hbar}\right)= \qquad (456)\\
&-T\left(\left\{\hat{s}_0\Phi(x)\cdot(\Phi^{\ddagger}(x),F)+(\Phi\leftrightarrow\Phi^{\ddagger})\right\}\otimes\Psi(y)\otimes e_{\otimes}^{iF/\hbar}\right)\\
&-T\left(\left\{(F,\Phi(x))\cdot(\Phi^{\ddagger}(x),F)+(\Phi\leftrightarrow\Phi^{\ddagger})\right\}\otimes\Psi(y)\otimes e_{\otimes}^{iF/\hbar}\right)\\
&+i\hbar T\left(\left\{\hat{s}_0\Phi(x)\cdot(\Phi^{\ddagger}(x),\Psi(y))+(F,\Phi(x))\cdot(\Phi^{\ddagger}(x),\Psi(y))+(\Phi\leftrightarrow\Phi^{\ddagger})\right\}\otimes e_{\otimes}^{iF/\hbar}\right)\quad \text{mod } \mathcal{J}_0,
\end{aligned}
$$

where again $F=\int(\lambda f\mathbf{L}_1+\lambda^2 f^2\mathbf{L}_2)$. One can now formulate a stronger set of local identities analogous to eq. (426), and one can prove these identities using T12c along the same lines as in the previous subsection, with $\mathbf{J}(y)$ there replaced everywhere by $\Psi(y)$. The potential anomaly of the stronger identities (and therefore the possible violation of eq. (456)) can now be removed by a suitable redefinition of the time ordered products $T_{n+2}(\mathbf{J}_0(x)\otimes\Psi_0(y)\otimes\mathbf{L}_1(x_1)\otimes\mathbf{L}_1(x_n))$ at n-th order in perturbation theory, where $\Psi=\Psi_0+\lambda\Psi_1+\lambda^2\Psi_2+\dots$. However, contrary to the case in the previous subsection, we now do not have to worry about potential symmetry issues, that had to be dealt with there, because Ψ_0 is always distinct from $\mathbf{J}_0$, the latter having ghost number 1.

4.9 Relation to other perturbative formulations of gauge invariance

In our approach to interacting quantum gauge theories, the gauge invarince of the theory was incorporated in the conditions that there exists a conserved interacting BRST-current operator, and that the corresponding charge operator be nilpotent. As we demonstrated, this follows from our Ward identity (322), the generating identity for T12a, T12b, and T12c. In the literature on perturbative quantum field theory in flat spacetime, other notions of gauge invariance of the quantum field theory have been suggested, and other conditions have been proposed to ensure those. We now briefly discuss some of these, and explain why these formulations are not suitable in curved spacetime.

Diagrammatic approaches (dimensional regularization): Historically, the first proofs of gauge invariance of the renormalized perturbation series in gauge theories on flat $\mathbb{R}^4$ were performed on the level of Feynman diagrams. The gauge-invariance of the classical Lagrangian implies certain formal identities between the diagrams at the unrenormalized level. At the renormalized level, these identities in turn would formally[15] imply the gauge-invariance of amplitudes. One must thus prove that these identities remain valid at the renormalized level. For this, it is important to have a regularization/renormalization scheme that preserves these identities. Such a scheme was found by 't Hooft and Veltmann [69, 70, 71], namely dimensional regularization. Because that scheme is also very handy for calculations (except for certain calculations involving Dirac-matrices), it has remained the most popular approach among practitioners. Modern presentations of this approach based on the Hopf-algebra structure behind renormalization in the BPHZ-approach [21, 22, 83] are [105, 106].

In curved space, scattering amplitudes are not well-defined, because there is no sharp notion of particle in general. At a more formal level, diagrammatic expansions in general are problematic because there does not exist a unique Feynman propagator, so a given Feynman diagram can mean very different mathematical expressions depending on one's choice of Feynman propagator. One may of course expand the theory using any Feynman propagator. However, then the problem arises that the Feynman propagator is not a local covariant functional of the met-

[15] We say "formally," because amplitudes can have additional infra-red divergences, which are very hard to treat in a gauge-invariant manner.

ric, but also depends upon boundary/initial conditions, which are intrinsically non-local. This would interfere with ones ability to reduce the ambiguity to local curvature terms. One might be tempted to take the local Feynman parametrix H_F, which is local and covariant. But this has the undesirable property that it is not a solution of the field equation, but only a Green's function modulo a smooth remainder, see Appendix D. This severely complicates the treatment of quantities that vanish due to field equations, and of the Ward identities. Finally, in curved space, the Feynman propagator is only well defined as a distribution in position space, while techniques such as dimensional regularization seem to work best in momentum spacetime. Thus, a diagrammatic proof of quantum gauge invariance of Yang-Mills theory in curved spacetime seems to be difficult and somewhat unnatural.

Zinn-Justin equation: In many formal approaches to perturbative gauge theory in flat spacetime $\mathbb{R}^4$, gauge invariance of the theory is expressed in terms of an integrated condition involving the so-called "effective action", $\Gamma_{\text{eff}}(S)$ of the theory associated with the classical action $S = S_0 + \lambda S_1 + \lambda^2 S_2$. The effective action is a generating functional for the 1-particle irreducible Feynman diagrams of the theory. The condition for perturbative gauge invariance is simply and elegantly encoded in the relation [112]

$$(\Gamma_{\text{eff}}(S), \Gamma_{\text{eff}}(S)) = 0\,. \tag{457}$$

Condition (457) is referred to as the "Slavnov Taylor identity" in "Zinn-Justin form". It is closely related to the "master equation" that arises in the Batalin-Vilkovisky formalism [6] (see also [60]), and it reduces to the classical condition $(S,S) = 0$ for BRST-invariance when one puts $\hbar = 0$. At the formal level, the Slavnov-Taylor identity is most straightforwardly derived from the path integral. It is also in this setting that one can understand relatively easily that it formally implies the absence of (infinite) counterterms to the classical action violating gauge invariance. However, by itself, it does not imply the gauge invariance of physical quantities such as scattering amplitudes, or identities like $Q_I^2 = 0$.

The effective action $\Gamma_{\text{eff}}(S)$ is only a formal quantity, since it involves integrations over all of spacetime. These integrations typically lead to infra-red divergences, as is in particular the case also in pure Yang-Mills theory. Therefore, also the Slavnov-Taylor equation (457) is only a formal identity. If the interaction $\lambda S_1 + \lambda^2 S_2$ is replaced by a local interaction, $F = \int \{\lambda f \mathbf{L}_1 + \lambda^2 f^2 \mathbf{L}_2\}$, with f a smooth cutoff function of compact support, then the infra-red divergences are avoided, and the effective action $\Gamma_{\text{eff}}(S_0 + F)$ is well defined. The precise definition of $\Gamma_{\text{eff}}(S_0 + F)$ within our framework is given in Appendix B. However, for the cutoff-interaction, the Slavnov-Taylor identity no longer holds. Nevertheless, it can be shown that $\Gamma_{\text{eff}}(S_0 + F)$ satisfies an analogous equation, given by eq. (491). That equation can be used to formally "derive" eq. (457), if one could prove that the anomaly in eq. (491) vanishes. Since the anomaly is closely related to the failure of the interacting BRST-current to be conserved, one might expect to be able to remove the anomaly by an argument similar to our proof of T12a, but this has not been worked out even in flat spacetime.

In curved spacetime, we may still define an effective action, $\Gamma_{\text{eff}}(S_0 + F)$, which now depends upon the arbitrary choice of a quasifree Hadamard state ω, see Appendix E. Hence it is

definitely not a quantity that depends locally and covariantly upon the metric, but also on the non-local choice of ω, Therefore, even at the formal level, it is not clear that the Slavnov-Taylor identity can be viewed as a renormalization condition that is compatible with the locality and covariance of the time-ordered products. Also, while the Slavnov-Taylor identity can again be formally derived from our Ward-Identity T12a, it does not directly imply the gauge-invariance of physical quantities such as n-point functions, and it also does not prove (even formally) that the OPE closes among physical operators. For these reasons, we prefer to work with the Ward-identities T12a, T12b, T12c in this paper, which are rigorous, and have a local and covariant character. Despite the above differences, the Zinn-Justin is probably to be regarded as the closest analogue to our renormalization conditions expressing local gauge invariance. The similarities can be made more explicit using our generating formula (322) [or eq. (336)] for our Ward identities.

Causal approach: A condition expressing perturbative gauge invariance in flat spacetime that is of a more local nature than (457) has been proposed in a series of papers by Dütsch et al. [29, 31, 30, 32, 33, 96], see also [73, 74, 75, 56, 57, 58]. These works are also related to the "quantum Noether condition" [75]. Let $T_n(x_1,\dots,x_n)$ be the time-ordered product of $T_n(\mathbf{L}_1(x_1)\otimes\cdots\otimes\mathbf{L}_1(x_n))$. (in the above papers, the interaction Lagrangian 4-form is here identified with a scalar by taking the Hodge dual). Let Q_0 be the free BRST-charge. Then it is postulated that there exists a set of time-ordered products $T_{n/l}(x_1,\dots,x_n)$ with the insertion[16] of some (unspecified) 3-form-valued field in the l-th entry such that

$$[Q_0, T_n(x_1,\dots,x_n)] = i\hbar\sum_{l=1}^{n} d_l T_{n/l}(x_1,\dots,x_n) \quad \text{modulo } \mathcal{J}_0 \tag{458}$$

for all $n>0$, where $d_l = dx_l^\mu \wedge \partial/\partial x_l^\mu$ is the exterior derivative acting on the l-th entry. The condition is to be viewed as a normalization on the time ordered products involving n factors of the interaction $\mathbf{L}_1$. Note that there are no explicit[17] conditions imposed on time-ordered products involving $\mathbf{L}_2$. Note also that the condition is imposed only modulo $\mathcal{J}_0$, that is, on shell. In fact, the authors of the above papers always work in a representation, where the field equations automatically hold (see section 3), rather than at the algebraic level, where the field equations need not be imposed as a relation. A related difference is that the above authors do not work with anti-fields, without which it appears to be very cumbersome to obtain powerful consistency relations for potential anomalies of (458). (Some aspects of this difference are addressed in [4].)

The key motivation for condition (458) is that, as our condition T12a), it formally implies that the S-matrix commutes with Q_0 in the "adiabatic limit," see above. Indeed, if we formally integrate (458) over $(\mathbb{R}^4)^n$, then the right hand side formally vanishes, being a total derivative.

[16]Thus in particular, $T_{n/l}(x_1,\dots,x_n)$ should be symmetric in all variables except x_l, and it is a 3-form in x_l.

[17] As explained in the above papers, however, implicit normalization conditions on time ordered products with factors of $\mathbf{L}_2$ arise from (458). Also, (458) apparently may even be used to determine the form of $\mathbf{L}_1$, which is simply given in our approach.

This shows that $\mathcal{S}$ formally commutes with Q_0. However, unlike our Ward identities, we do not believe that eq. (458) would imply $Q_I^2 = 0$ for the interacting BRST-charge, or $[Q_I, \Psi_I] = 0$ for gauge invariant operators.

The relation (458) is apparently different from our corresponding condition T12a (considered in flat spacetime), so we now briefly outline how they are related. Consider a prescription for the time-ordered products satisfying our Ward identity T12a, so that, in particular, eq. (458) does not hold for that prescription. However, let us now make the following redefinition of the time-ordered products containing two factors of $\mathbf{L}_1$, that is,

$$T_2(\mathbf{L}_1(x_1) \otimes \mathbf{L}_1(x_2)) \to T_2(\mathbf{L}_1(x_1) \otimes \mathbf{L}_1(x_2)) + T_1(\mathbf{L}_2(x_1, x_2)), \tag{459}$$

where we recall the notation $\mathbf{L}_2(x_1, x_2) = 2\mathbf{L}_2(x_1)\delta(x_1, x_2)$. Let us further note that

$$\hat{s}_0 \mathbf{L}_2(x_1, x_2) + (\mathbf{L}_1(x_1), \mathbf{L}_1(x_2)) = d_1 \mathcal{O}_{2/1}(x_1, x_2) + d_2 \mathcal{O}_{2/2}(x_1, x_2) \tag{460}$$

for some fields $\mathcal{O}_{2/1} \in \mathbf{P}^{4/3}$ and $\mathcal{O}_{2/2} \in \mathbf{P}^{3/4}$ supported on the diagonal, and $\hat{s}_0 \mathbf{L}_1 = d\mathcal{O}_1$. Using that $[Q_0, T_n] = i\hbar \hat{s}_0 T_n$ modulo $\mathcal{J}_0$, and defining $T_{n/l}$ by

$$\begin{aligned} T_{n/l}(x_1, \dots, x_n) &= \sum_{j=1,2} T_{n-1}\big(\mathbf{L}_1(x_1) \otimes \dots \mathcal{O}_{2/j}(x_{l+j-1}, x_{l+j}) \otimes \dots \mathbf{L}_1(x_n)\big) + \text{cycl. perm.} \\ &+ T_n\big(\mathbf{L}_1(x_1) \otimes \dots \mathcal{O}_1(x_l) \otimes \dots \mathbf{L}_1(x_n)\big), \end{aligned} \tag{461}$$

one can then check that eq. (458) holds. Thus, our Ward identity implies (458) if a finite renormalization change is made, and presumably (458) may also be used to deduce our Ward identity T12a. Note, however, that our identities T12b and T12c are conditions that go definitely beyond the Ward-identities (458).

5 Summary and outlook

In this paper, we have given, for the first time, a perturbative construction of non-abelian Yang-Mills theory on arbitrary globally hyperbolic curved, Lorentzian spacetime manifolds. Following earlier work on quantum field theory in curved spacetime, our strategy was to construct the interacting field operators and the algebra that they generate. This was accomplished starting from a gauge fixed version of the theory with ghost and anti-fields, and then defining the algebra of observables of perturbative Yang-Mills theory as the BRST-cohomology of the corresponding algebra associated with the gauge fixed theory. To implement this strategy it was necessary to first find a prescription for defining a *conserved* interacting BRST-current, and for which the corresponding conserved charge is furthermore *nilpotent*. We were able to characterize such a prescription by a novel set of Ward identities for the time-ordered products in the underlying free theory. We furthermore showed how to find a renormalization prescription for which the Ward-identities indeed hold. In addition, we showed that our renormalization prescription also

satisfies other other important properties, notably the condition of general covariance. Altogether, these constructions provide a proof that perturbative Yang-Mills theory can be defined as a consistent, local covariant quantum field theory (to all orders in perturbation theory), for any globally hyperbolic spacetime.

A key feature of our approach is that it is entirely local in nature, in the sense that our renormalization conditions only make reference to local quantities. A local approach is essential in a generic curved spacetime in order find the correct renormalization prescription respecting locality and general covariance. But it is also advantageous in flat spacetime in many respects compared to other existing approaches in flat spacetime, such as approaches focused on the scattering matrix, or approaches based on the path-integral. The key advantages of our approach are the following:

- Because our approach is completely local, we can completely disentangle the the infra-red divergences and ultra-violet divergences of the theory. This is mandatory in Yang-Mills theory, where infra-red divergences pose a major problem, even in flat spacetime.

- Because our approach is algebraic in nature, the objects of primary interest are the interacting field operators, rather than auxiliary quantities such as effective actions or scattering matrices. This makes it easy for us to prove the important result that the operator product expansion of Yang-Mills theory closes among gauge invariant fields, and that the renormalization group flow does not leave the space of gauge invariant fields. On the other hand, it tends to be much more complicated to prove such statements in other formalisms even in flat spacetime.

- Because our approach is local and covariant, we can directly analyze the dependence of our constructions on the metric. For example, one can directly obtain the following result: If a non-abelian gauge theory has trivial RG-flow in flat spacetime (such as the $\mathrm{N} = 4$ super Yang-Mills theory), then it also must have trivial RG-flow in any spacetime in which possible renormalizable curvature couplings in the Lagrangian (such as a $R\,\mathrm{Tr}\Phi^2$-type term) happen to vanish. Thus, the $\mathrm{N} = 4$ super Yang-Mills theory has trivial RG-flow in any spacetime with vanishing scalar curvature. Note that, unlike in flat spacetime, this does by no means imply that the theory is conformally invariant, because a spacetime with $R = 0$ will not in general admit any conformal isometries.

A weak point of our constructions, as for most other perturbative constructions in quantum field theory, is that one does not have any control over the convergence of the perturbation series. This is in particular a problem for quantum states such as bound states that are not expected to have a perturbative description. A partial resolution of this problem is provided by the operator product expansion (see sec. 4.2), because it allows one to compute n-point correlation functions in terms of OPE-coefficients and 1-point functions (“form factors”), which one may regard as additional phenomenological input. But a full solution would presumably require to go beyond perturbation theory, which seems a distant goal even in flat spacetime.

Apart from this problem, there remain a couple of technical questions related to the perturbation expansion, of which we list a few:

5.1 Matter fields, anomalies

In this paper, we have considered only pure Yang-Mills theory for simplicity. Clearly, one would like to add matter fields, such as fermion fields in a representation R of the gauge group G. In that case, the general strategy and methods of our paper can still be applied. But it is no longer clear that the Ward-identities formulated in this paper can still be satisfied, as there can now be non-trivial solutions to the corresponding consistency conditions in the presence of chiral fermions. If the Ward-identities cannot be satisfied, one speaks of an anomaly. In our case this would imply that the interacting BRST-current is no longer conserved, and that a conserved BRST-charge cannot be defined, meaning that the theory is inconsistent at the quantum level. In flat space, this can happen if the gauge group contains factors of $U(1)$, for certain representations R. By the general covariance of our construction, the types of anomalies in flat space must then also be absent in any curved spacetime. However, in curved space, a new type of anomaly can also arise in the presence of chiral fermions and abelian factors in the gauge group. For example, even at the level of free Yang-Mills theory, one can compute that the divergence $d\mathbf{J}_I$ (exterior differential) of the quantum BRST current operator is not zero as required by consistency, but it has a contribution to its divergence proportional of the type given in eq. (63), which cannot be eliminated by finite renormalization. In particular, one finds a contribution $d\mathbf{J}_I \propto \mathcal{A}_I + \dots$ at 1-loop order, where

$$\mathcal{A} = \text{const.} \sum_K \text{Tr}[\text{R}(T_K)]\, C^K\, \text{Tr}(R \wedge R) \tag{462}$$

and where the sum over K is over the abelian generators of the Lie-algebra only. In the standard model, with gauge group $G = SU(3) \times SU(2) \times U(1)$, the representation of the abelian generator Y (charge assignments of the fermion fields) is precisely so that $\mathcal{A} = 0$, as also observed by [54, 88]. However, we do not know whether the theory remains free of this kind of anomaly to arbitrary orders in the perturbation series. This would be important to check.

It is also important to investigate whether the renormalization conditions considered in this paper can be used to show that a divergence-free interacting stress tensor $T_I^{\mu\nu}$ can be constructed. Here, one can presumably use the techniques of [66] to show that there is no anomaly for this conservation equation, but it would be important to settle the details. A particularly interesting question in this connection is to see precisely how the expected trace anomaly for this quantity arises in the present framework.

5.2 Other gauge fixing conditions

In this paper, we have worked with a specific gauge fixing condition (the Lorentz gauge). The important feature of this condition for our purposes was that the field equation for the spin-1

field then becomes $\Box A + \cdots = 0$, where the dots represent terms with less derivatives. This was important because only in that case are we able to construct a Hadamard parametrix for the vector field, which is a key ingredient in our constructions. However, one may wish to consider other types of gauge fixing conditions, both for practical purposes, as well as a matter of principle. Even if a Hadamard parametrix could still be defined in such cases, it is not a priori clear that the theories defined using different gauge fixing conditions are equivalent. In our approach, equivalence would mean that the algebras of observables obtained from different gauge fixing conditions are canonically isomorphic. We have not investigated the question whether this is indeed the case.

5.3 Background independence

In our constructions (as in all other standard approaches to perturbative Yang-Mills theory), we have split the Yang-Mills connection $\mathcal{D} = \nabla + i\lambda A$ into the standard flat, non-dynamical background connection ∇, and a dynamical field A. At the level of classical Yang-Mills theory it is evident that it is immaterial how this split is made, i.e., classical Yang-Mills theory is background independent in this sense. In particular, the standard choice $\nabla = \partial$ in flat spacetime is just one possibility among infinitely many other ones. In the gauge fixed classical theory with ghosts and anti-fields, different choices of the background connection give rise to different classical actions. The difference is, however, only by a BRST-exact term. Since the classical theory is defined as the BRST-cohomology, such a BRST-exact term does not change the brackets between the physical observables, and hence the theory is background independent also in the gauge-fixed formalism. Unfortunately, we do not know whether the same statement is still true in the quantum field theory, i.e., we do not know whether the algebras of physical observables associated with different choices of the background connection are still isomorphic. The difficulty is that, in quantum field theory, the background connection ∇ is treated very differently from the dynamical part A: The background connection would enter the definition of the propagators, e.g., of the local Hadamard parametrices, while A is a quantum field.

The question whether one is allowed to shift parts of A into ∇ and vice versa is closely related to the question whether the "principle of perturbative agreement" formulated in [66] can be satisfied with respect to the gauge connection. The satisfaction of this principle is equivalent to certain Ward-identities at the level of the time-ordered products, but we do not know in the present case whether these Ward identities can be satisfied, i.e., whether there are any anomalies. In [66], a potential violation of these identities may be identified with a certain cohomology class. In our case, when the background structure in question is a gauge connection, the potential violation would be represented by a certain 2-cocycle on the space of all gauge potentials. An anomaly of this sort could arise in theories with chiral fermions. Thus, the question of background independence in quantum Yang-Mills theory remains an open problem, which has not been solved, to our knowledge, even in flat spacetime.

Acknowledgments: I would like to thank F. Brennecke, D. Buchholz, M. Dütsch, L. Faddeev, K. Fredenhagen, D. Grigore and R. M. Wald for discussions, and I would also like to thank M. Henneaux for discussions at an early stage of this work. Some parts of this work were completed during the 2007 program "Mathematical and Physical Aspects of Perturbative Approaches to Quantum Field Theory" at the Erwin-Schroedinger Institute, Vienna, to which I express my gratitude for its financial support and hospitality.

A U(1)-gauge theory without vector potential

In the case of a pure $U(1)$-gauge theory, one may consider a different starting point for defining the theory, using as the basic input only the field equations for the 2-form field strength tensor rather than the action for the gauge potential A. This is because the field equations may then be written without reference to the gauge potential as equations for the field strength F, viewed now as the dynamical variable. The equations are of course Maxwell's equations, in differential forms notation $dF = 0$ and $d * F = 0$.

On a curved manifold M with nontrivial topology, not every closed form F need to be exact, so it does not follow from the field equation $dF = 0$ that F can be written in terms of a vector potential as $F = dA$. Thus, using only Maxwell's equations as the input defines a more general theory classically than the action $\int dA \wedge * dA$, because cohomologically non-trivial solutions F are possible. In this section, we briefly indicate how one may quantize such a theory.

A globally hyperbolic spacetime always has topology $M = \Sigma \times \mathbb{R}$, so closed but non-exact 2-forms F can exist on M if Σ contains any non-contractible 2-cycles, C. Let us cover M by

$$M = \bigcup_i M_i \tag{463}$$

where each M_i a globally hyperbolic, connected and simply connected spacetime in its own right, which does not contain any non-contractible 2-cycles. Consequently on each M_i, any closed 2-form is exact, and the classical theory defined by Maxwell's equations $dF = 0$, $d * F$ is completely equivalent to the theory of a vector potential A with action (34). Thus, by the results of the previous sections, we can construct a corresponding algebra of observables $\hat{\mathcal{F}}_0(M_i)$ for each i, containing gauge-invariant observables such as polynomials of the field strength.

Each $\hat{\mathcal{F}}_0(M_i)$ is only given to us as an abstract *-algebra, so we do not a priori know what is the relation between those algebras for different i. However, if M_i is contained in M_j, then by the general covariance property, there is an embedding of algebras $\alpha_{i,j} \equiv \alpha_{\psi(i,j)} : \hat{\mathcal{F}}_0(M_i) \to \hat{\mathcal{F}}_0(M_j)$, where $\psi(i,j) : M_i \to M_j$ is the embedding. Thus, following ideas of Fredenhagen, and Küskü [47, 48, 85], we may define an algebra $\mathcal{A}_u(M)$ as the universal algebra

$$\mathcal{A}_u(M) \equiv \text{ind}-\lim_{M_i} \hat{\mathcal{F}}_0(M_i)\,. \tag{464}$$

The universal algebra is defined as the unique algebra such that there exist *-homorphisms $\alpha_i : \hat{\mathcal{F}}_0(M_i) \to \mathcal{A}_u(M)$ with the property $\alpha_j \circ \alpha_{j,i} = \alpha_i$. It is characterized by the fact there

are no additional relations in $\mathcal{A}_u(M)$ apart from the ones in the subalgebras. Thus, $\mathcal{A}_u(M)$ is generated by the symbols $F_i(f)$ where supp $f \subset M_i$, which we think of as smeared field strength tensors

$$F_i(f) = \int_{M_i} f \wedge F \,. \tag{465}$$

Their relations are

$$F_i(f) = F_j(f), \quad \text{if supp } f \subset M_i \cap M_j, \tag{466}$$

and the $F_i(f)$, with supp $f \subset M_i$ satisfy all the relations in $\hat{\mathcal{F}}_0(M_i)$, which are

$$[F_i(f), F_i(h)] = i\Delta(f,h)\,\mathbb{1}\,, \quad F_i(df) = 0 = F_i(*df)\,, \tag{467}$$

for any 1-forms f,h of compact support in M_i. Here, $\Delta : \Omega^2_0(M) \times \Omega^2_0(M) \to \mathbb{R}$ denotes the advanced minus retarded fundamental solution for the hyperbolic operator $\delta d + d\delta$ acting on 2-forms.

For an arbitrary compactly supported 2-form f on M, we may then define the algebra element $F(f) \in \mathcal{A}_u(M)$ as

$$F(f) \equiv \sum_i F_i(\psi_i f)\,, \tag{468}$$

where supp $\psi_i \subset M_i$, and $\sum_i \psi_i = 1$ on supp f. It is not difficult to show using eq. (466) that this definition does not depend upon the particular choice of the covering. From eq. (467), it then also follows that $F(df) = 0 = F(d^* f)$ holds for arbitrary compactly supported forms f in M. One can also easily show that $F(f) \star_\hbar F(h) - F(h) \star_\hbar F(f) = 0$ for any two test-forms having spacelike related support. Indeed, after splitting f,h using a suitable a partition of unity, we may assume that the supports of f and h are contained in sets M_i and M_j. Since M is assumed to be connected, there exists therefore a globally hyperbolic spacetime $N \subset M_i \cup M_j$ in which every 2-cycle is contractible, and we may assume that N appears in the covering of M. We may then view both $F(f)$ and $F(h)$ as elements in $\hat{\mathcal{F}}_0(N)$, where they commute. Since Δ is uniquely determined by its action on testfunctions supported in a neighborhood of a Cauchy surface, it then also follows that $[F(f), F(h)] = i\hbar\Delta(f,h)\,\mathbb{1}$.

The universal algebra contains certain central elements that carry information about the topology of M. They arise as follows. Let C be a 2-cycle in M, and let $\{\psi_i\}$ be a partition of unity subordinate to the covering $\{M_i\}$ of M. By Poincare duality, we can find a closed 1-form h_C on M such that

$$\int_M h_C \wedge \alpha = \int_C \alpha \tag{469}$$

for any closed 2-form α, and we may arrange h_C to have support in a neighborhood of C. The 2-form $\psi_i h_C$ has compact support in M_i, and we may define

$$Z_e[C] = F(h_C) \equiv \sum_i F_i(\psi_i h_C) \in \mathcal{A}_u(M)\,. \tag{470}$$

We claim that $Z_e[C]$ is independent of the particular choice of h_C, and of the partition $\{U_i, \psi_i\}$. Independence of the partition was already shown above for general 2-forms. To show independence of h_C, consider another h'_C with the same properties, and let $h_C - h'_C = \omega$. Then ω is closed, of compact support and, $\int \omega \wedge \alpha = 0$ for any closed 2-form α. By the well-known fact that the pairing

$$\int : H^2(M) \otimes H^2_0(M) \to \mathbb{R} \tag{471}$$

is non-degenerate, we therefore must have that $[\omega] = 0$ in $H^2_0(M)$, i.e., $\omega = d\beta$ for some 1-form β of compact support. Independence of $Z_e[C]$ on the particular form of h_C then follows from $F(d\beta) = 0$.

It then also follows that $Z_e[C]$ only depends upon the homotopy class of C, i.e., $Z_e[C]$ may be viewed as a map

$$Z_e : H_2(M;\mathbb{Z}) \to \mathcal{A}_u(M), \quad [C] \mapsto Z_e[C]\,. \tag{472}$$

In particular $Z_e[C] = 0$ for any 2-cycle C that can be deformed into a point. Because $Z_e[C]$ only depends upon the class $[C]$ of C in $H_2(M)$, it follows that, given any sufficiently small compact region $K \subset M$, we may deform C so as to be in the causal complement of K, that is $C \subset J^+(K) \cup J^-(K)$. By choosing h_C to be supported in a sufficiently small neighborhood of C, it then follows that

$$[Z_e[C], F(f)] = 0, \quad \forall f \in \Omega^2_0(K)\,, \tag{473}$$

But then this also holds for arbitrary f of compact support, because f may be written as $\sum \psi_i f$, with each $\operatorname{supp}\psi_i$ so small that C and hence $\operatorname{supp} h_C$ can be deformed so as to lie in the causal complement. Thus, $Z_e[C]$ is in the center $\mathcal{Z}(\mathcal{A}_u(M))$ of $\mathcal{A}_u(M)$. By taking the dual of h_C in eq. (470), we may similarly define

$$Z_m[C] = \sum_i F_i(\psi_i * h_C) \in \mathcal{Z}(\mathcal{A}_u(M))\,, \tag{474}$$

and this quantity has similar properties as $Z_e[C]$.

The center-valued quantities $Z_e[C], Z_m[C]$ correspond to the electric and magnetic fluxes through a 2-cycle C. They are analogous to the classical quantities $\int_C F$ respectively $\int_C *F$ and satisfy the same additivity relations under the addition of cycles. Other interesting derived quantities may also be defined. For example, let $C_1, C_2, \dots$ be a basis of 2-cycles in $H_2(M;\mathbb{Z})$, and let

$$(Q^{-1})_{jk} = I(C_j, C_k) \tag{475}$$

be the matrix of their intersection numbers. Then we may define

$$q_{top} = \sum_{j,k}^{b_2} Q^{jk} Z_e[C_i] Z_e[C_k] \quad \in \mathcal{Z}(\mathcal{A}_u(M)) \tag{476}$$

and this is analogous to the classical topological quantity

$$q_{class} = \int_M F \wedge F = \sum_{j,k} Q^{jk} \Big(\int_{C_j} F \Big) \Big(\int_{C_k} F \Big) \tag{477}$$

by the so-called “Riemann identity” for closed differential forms.

In any factorial representation $\pi : \mathcal{A}_u(M) \to \mathrm{End}(\mathcal{H})$ on a Hilbert space $\mathcal{H}$, the representers corresponding to $Z_e[C], Z_m[C]$ are by definition represented by multiples of the identity, i.e.,

$$\pi(Z_e[C]) = c_e[C] \cdot I, \quad \pi(Z_m[C]) = c_m[C] \cdot I. \tag{478}$$

where c_e, c_m are valued in the complex numbers. By DeRahm’s theorem, they can be represented by 2-forms f_e and f_m, both of which must be closed. Choosing a basis $\{\omega^i\}$ of $H^2(M)$, for example dual to a basis of 2-cycles $\{C_i\}$, we may thus expand $f_e = \sum_i q_i \omega^i$, and $f_m = \sum_i g_i \omega^i$ with numerical constants $q_i, g_i \in \mathbb{R}$ depending upon the representation. These constants are then the (canonically normalized) numerical values of the electric and magnetic flux through the respective cycle in the representation π.

The above construction of Maxwell theory (without a vector potential) is somewhat abstract, and we now discuss an equivalent description. As above, let $\{\omega^i\}$ be a set of closed forms forming a basis of $H^2(M)$. Any closed form F may thus be written uniquely as $F = dA + \sum_i q_i \omega^i$. Substitution into the action S gives

$$S = \frac{1}{2} \int dA \wedge *dA + j \wedge *A \tag{479}$$

where $j = \sum q_i \delta\omega^i$ is considered as an external (conserved) current coupled to A. The quantization of this theory now proceeds along similar lines as for the action S without the external current. We correspondingly get an algebra of observables $\mathcal{A}_q(M)$, which now depends upon the choice of $q \equiv \{q_i\}$ and $\{\omega^i\}$ through the external current. The algebra is spanned by generators $\int f \wedge dA$, and

$$\widehat{F}(f) = \int f \wedge dA + \sum q_i \Big(\int \omega^i \wedge f \Big) \mathbb{1}. \tag{480}$$

They satisfy the same relations as the generators $F(f)$ above in the algebra $\mathcal{A}_u(M)$. From this it may be seen that the algebra $\mathcal{A}_q(M)$ only depends upon q_i and the equivalence classes $[\omega^i]$. This algebra also has further relations not present in $\mathcal{A}_u(M)$, because the elements $\widehat{Z}_e[C] \in \mathcal{A}_q(M)$ defined in the same way as the central elements $Z_e[C] \in \mathcal{A}_u(M)$ above, are now represented by multiples of the identity, namely

$$\widehat{Z}_e[C] = \sum q_i \Big(\int_C \omega^i \Big) \mathbb{1} \quad \in \mathcal{A}_q(M), \tag{481}$$

while the elements $Z_e[C] \in \mathcal{A}_u(M)$ are only in the center, but not necessarily proportional to the identity. Thus, $\mathcal{A}_u(M)$ and $\mathcal{A}_q(M)$ are not isomorphic. Instead, we have

$$\mathcal{A}_u(M) \cong \int^{\oplus} \prod_{i=1}^{b_2} dq_i \, \mathcal{A}_q(M). \tag{482}$$

By contrast, the magnetic fluxes $\widehat{Z}_m[C]$, defined as above, are not proportional to the identity but only elements in the center of $\mathcal{A}_q(M)$. This apparent asymmetry between the electric and

magnetic fluxes arises from the fact that we have chosen to quantize the theory starting from a potential for F, rather $*F$, which would also be possible. Then the roles of electric and magnetic fluxes would be reversed.

A physically relevant example of a spacetime M with a non-trivial 2-cycle is the Kruskal extension of the Schwarzschild spacetime. It has line element

$$ds^2 = \frac{32M^3 e^{r/2M}}{r}(-dT^2 + dX^2) + r^2(d\theta^2 + \sin^2\theta\, d\varphi^2) \quad r > 0\,, \tag{483}$$

and topology $M = \mathbb{R}\times\mathbb{R}\times S^2$, where r is defined through $T^2 - X^2 = (1-r/2M)e^{r/2M}$. It is a globally hyperbolic spacetime with a non-trivial 2-cycle, homotopic to S^2. Hence, the universal algebra possesses non-trivial central elements $Z_e[S^2], Z_m[S^2]$, and this gives rise to the possibility of having non-trivial electric and magnetic fluxes in that spacetime, as also realized by Ashtekar et al. [3].

We now sketch an argument that arbitrary values of the electric and magnetic charges may be realized in representations π carrying a unitary representation of the time-translation symmetry group. The spacetime is a solution to the vacuum Einstein-equation $R_{\mu\nu} = 0$, with static timelike Killing field $K = \partial/\partial t$, with $t = 4M\tanh^{-1}(X/T)$. By the standard identity $\nabla_{[\mu}(\varepsilon_{\nu\sigma]\alpha\beta}\nabla^\alpha K^\beta) = \frac{2}{3}R_{\alpha\beta}K^\beta e^\alpha{}_{\mu\nu\sigma}$ valid for any Killing field K, $\phi_{\mu\nu} = \frac{1}{4\pi}\nabla_{[\mu}K_{\nu]}$ is therefore a static (meaning $\pounds_K\phi = 0$) solution to the classical Maxwell equations. Given $q,g\in\mathbb{R}$, we define $\gamma_{p,q} : F(f)\mapsto F(f) + q\int_{S_2} f\wedge\phi\,\mathbb{1} + g\int_{S_2} f\wedge *\phi\,\mathbb{1}$. This is an automorphism of $\mathcal{A}_u(M)$. Let us assume that there is a factorial vacuum state $\langle\,.\,\rangle_0$ on $\mathcal{A}_u(M)$ invariant under the action of the time-translation isometries (which can presumably be constructed by the techniques of Junker et al. [78]), and let us assume that $\langle Z_e[S^2]\rangle_0 = 0 = \langle Z_m[S^2]\rangle_0$. Then the states $\langle\,.\,\rangle_{q,g} = \langle\gamma_{q,g}(\,.\,)\rangle_0$ are also factorial and the corresponding GNS-representation carry a unitary representation of the time-translation symmetries, with invariant vacuum vector. Furthermore, by $\int_{S^2} *\phi = 1$, we have

$$\pi_{q,g}(Z_e[S^2]) = qI, \quad \pi_{q,g}(Z_m[S^2]) = gI\,. \tag{484}$$

in the corresponding GNS-representations $\pi_{q,g}$ of these states. Thus, the representations $\pi_{q,g}$ carry electric flux q and magnetic flux g. In this sense, the numbers q,g may be viewed as superselection charges, as also noted by Ashtekar et al. [3].

B Effective Actions in curved spacetime

We here give the definition of the effective action in our framework following [14, 15] and a derivation of a set of consistency conditions. We also emphasize that the effective action is a state dependent quantity, and therefore, unlike the T-products, does not have a local, covariant dependence upon the metric.

In the path integral formulation of quantum field theory, the effective action in a scalar field theory is formally defined as follows (see e.g., [110]). Let $j \in C_0^\infty(M)$ be an external current density, and define, formally,

$$\exp(Z^c(j)) = \int [\mathcal{D}\phi] \exp\left(iS/\hbar + \int j\phi\right). \tag{485}$$

Then the effective action $\Gamma_{\rm eff}$ is defined, again formally, as the Legendre transformation of $Z^c(j)$: Define ϕ through $\phi = \delta Z^c(j)/\delta j$, and $\Gamma_{\rm eff} = \int j\phi - Z^c(j)$. The quantity $\Gamma_{\rm eff}$ is a formal power series in $\hbar$ depending on ϕ (and the action S), and may thus be viewed as an element of $\mathcal{F}$. The above construction is formal in several ways: The quantity $Z^c(j)$ is typically viewed as the generating functional for the hierarchy of connected time-ordered n-point functions of the quantum field ϕ. It thus depends upon a choice of state, and the same is consequently true for the effective action. This is obscured in the above functional integral formulation. Here, the choice of state would enter the precise choice of the formal path-integral measure $[\mathcal{D}\phi]$. Also, because the path-integral derivation does not specify the precise definition of the path-integral measure $[\mathcal{D}\phi]$, it necessarily disregards all issues related to renormalization. We therefore now give a precise definition of the effective action in curved spacetime.

For this, we define, following [14], the quantities $T^c_\omega : A^{\otimes n} \to \mathbf{W}_0$ (A the space of local actions) implicitly by

$$T(\exp_\otimes(iF/\hbar)) = \sum_{n \geq 0} \frac{1}{n!} : T^c_\omega(\exp_\otimes iF/\hbar) \cdots T^c_\omega(\exp_\otimes iF/\hbar) :_\omega, \tag{486}$$

where the n-th term has n factors. Unlike T, the quantity T^c_ω is not local and covariant, but depends upon the global choice of ω. It can be shown that $\tau^c_\omega(F^{\otimes n}) = \lim_{\hbar \to 0} T^c_\omega(F^{\otimes n})/\hbar^{n-1} \in A$ exist. Next, define a functional $\Gamma_\omega : A^{\otimes n} \to \mathbf{W}_0$ implicitly by

$$\tau^c_\omega\left(\mathrm{e}_\otimes^{i\Gamma_\omega(\exp_\otimes F)/\hbar}\right) = T^c_\omega\left(\mathrm{e}_\otimes^{iF/\hbar}\right). \tag{487}$$

It can be shown that, for $F \in A$

$$\Gamma_\omega(1\!\!1) = 0, \quad \Gamma_\omega(F) = F, \tag{488}$$

as well as

$$\Gamma_\omega(\mathrm{e}_\otimes^F) = F + O(\hbar). \tag{489}$$

Given an interaction $F \in A$, we define an "effective action" (with respect to the state ω) associated with $S_0 + F$ by

$$\Gamma_{\rm eff}(S_0 + F) = S_0 + \Gamma_\omega(\mathrm{e}_\otimes^F) = S_0 + F + O(\hbar), \tag{490}$$

Again, the higher order terms in $\hbar$ depend upon the state ω, and are not local and covariant. This property makes the effective action in general unsuitable to solve the renormalization problem

in curved spacetime, since the local and covariance properties of the renormalization procedure cannot be controlled.

The effective action obeys a useful identity that can presumably be used to analyze potential anomalies in the Ward identities (as an alternative to our approach), at least in flat spacetime. To formulate this identity, consider any local field polynomial $\mathcal{O}$, and the modified action $S_0 + F \to S_0 + F + \int_M h \wedge \mathcal{O}$, where $h \in \Omega_0(M)$ is a compactly supported smooth form. Then we have the identity [14]

$$\begin{aligned} &\int_M \frac{\delta \Gamma_{\text{eff}}(S_0 + F + \langle h, \mathcal{O} \rangle)}{\delta h(x)} \wedge \frac{\delta \Gamma_{\text{eff}}(S_0 + F + \langle h, \mathcal{O} \rangle)}{\delta \phi(x)} \Bigg|_{h=0} \\ = \quad &\int_M \frac{\delta}{\delta h(x)} \Gamma_{\text{eff}} \Big(S_0 + F + \langle h, \mathcal{O} \delta(S_0 + F)/\delta\phi \rangle + \Delta_{\mathcal{O}} \rangle \Big) \Bigg|_{h=0}, \end{aligned} \tag{491}$$

where $\Delta_{\mathcal{O}}(x) = \Delta_{\mathcal{O}}(\mathrm{e}_{\otimes}^{F})(x) \in A$ is the anomaly corresponding to $\mathcal{O}$ in the corresponding anomalous "Master Ward Identity" in sec. 4.4., see also [14, 15]. It is viewed here as a 4-form.

C Wave front set and scaling degree

We here recall the basic definition of the wave front set of a distribution and some of its elementary properties. For details, see [72]. If u is a compactly supported *smooth* function on $\mathbb{R}^n$, then by standard theorems of distribution theory, its Fourier transform, $\hat{u}(p) = (2\pi)^{-n/2} u(\exp(ip\,.))$ is an analytic function on $\mathbb{R}^n$ falling off faster than any inverse power of p, i.e.,

$$|\hat{u}(tp)| \le c_N (1 + |t|)^{-N}, \quad t \in \mathbb{R} \tag{492}$$

for some c_N not depending upon p, and any N. Conversely, this bound implies that a compactly supported distribution u is in fact smooth. The idea of the wave front set is to use the possible failure of this bound to characterize the non-smoothness of a distribution. For compactly supported *distributions* u, one defines the set of singular directions by

$$\Sigma(u) = \{ p \in \mathbb{R}^n \setminus 0 \mid |\hat{u}(tp) \ge c_N (1 + |t|)^{-N} \quad \text{for some } N, \text{ all } t > 0 \}. \tag{493}$$

We define the wave front set of any distribution at a point $x \in \mathbb{R}^n$ by

$$\mathrm{WF}_x(u) = \bigcap_{\psi : x \in \mathrm{supp}\, \psi} \Sigma(\psi u). \tag{494}$$

where the intersection is over all smooth compactly supported cutoff functions ψ. The wave front set is clearly invariant under dilatation, and therefore a cone, and it only depends on the behavior of u in an arbitrary small neighborhood of x. For distributions u defined on a smooth n-dimensional manifold X one defines the wave front set as follows. Let κ, U be a coordinate chart covering x. Then, choosing a smooth cutoff function Ψ supported in U that is 1 near x, we

can define $\kappa^*(\psi u)$, which is now a distribution that is defined on $\mathbb{R}^n$. We define the wave front set to be the set

$$\mathrm{WF}_x(u) = (\kappa^{-1})^* \mathrm{WF}_{\kappa(x)}(\kappa^*(\psi u)) \subset T^*_x X \tag{495}$$

It can be proved that this definition does not depend upon the arbitrary choice of κ, ψ, and one defines $\mathrm{WF}(u)$ to be the union of all $\mathrm{WF}_x(u)$. One relevant application of the wave front set in perturbative quantum field theory is the following theorem [72] about the product of distributions.

Theorem 5. Let u, v be distributions on X. If $0 \notin \mathrm{WF}_x(u) + \mathrm{WF}_x(v)$, then the pointwise product uv is defined in some neighborhood of x, and $\mathrm{WF}_x(uv) \subset \mathrm{WF}_x(u) + \mathrm{WF}_x(v)$.

Clearly, if the assumption holds for all $x \in X$, then the pointwise product is globally defined on X. Another useful theorem about wave front sets is the following [72]. Let $K \subset \mathbb{R}^n$ be a convex open cone, and let $u(x+iy)$ be analytic in $\mathbb{R}^n + iK$ for $|y| < \delta$ and some δ, with the property that $|u(x+iy)| \le C|y|^{-N}$ for some N, and all $y \in K$ with $|y| < \delta$. Then the boundary value $u(x) = \mathrm{B.V.}_{y \to 0}\, u(x+iy)$, with the limit taken for $y \in K$ defines a distribution on $\mathbb{R}^n$.

Theorem 6. The wave front set of $u(x) = \mathrm{B.V.}_{y \to 0}\, u(x+iy)$ with the limit taken within the cone K, i.e., $y \in K$, is bounded by

$$\mathrm{WF}(u) \subset \mathbb{R}^n \times K^D, \tag{496}$$

where $K^D = \{k \in \mathbb{R}^{n*} \mid\ k \cdot y < 0\ \forall y \in K\}$ is the dual cone.

In applications, one often deals with distributions that are solutions to a partial differential equation $Au = 0$, where A is partial differential operator on X (or even a pseudo-differential operator), i.e.,

$$A = \sum_{n=0}^{N} a^{\mu_1 \cdots \mu_n}(x) \nabla_{(\mu_1} \dots \nabla_{\mu_n)}\,. \tag{497}$$

Under this condition, it can be shown that the wave front set of u must be restricted to the set

$$\mathrm{WF}(u) \subset \{(x,k) \mid a^{\mu_1 \cdots \mu_N}(x) k_{\mu_1} \dots k_{\mu_N} = 0\}\,. \tag{498}$$

In case when A is the wave operator on a Lorentzian manifold, we hence learn that any distributional solution u of the wave equation can only have vectors of the form (x,k) in the wave front set when k is a null-vector. Another important application of the wave front set for quantum field theory in curved spacetime is the propagation of singularities theorem. Consider a distribution u on a spacetime (M,g) that is a solution to the wave equation $\Box u = f$, with f a *smooth* source. The wave operator defines a 1-particle Hamiltonian on "phase space" T^*M by $h(x,p) = g^{\mu\nu}(x) p_\mu p_\nu$, and Hamilton's equations, defined with respect to the symplectic structure $dx^\mu \wedge dp_\mu$,

$$\dot{p}_\mu = -2\Gamma_{\nu\mu\rho}(x) p^\nu p^\rho \tag{499}$$

$$\dot{x}^\mu = 2g^{\mu\nu}(x) p_\nu \tag{500}$$

define a flow in phase space, $t \mapsto \phi_t$, which is just the geodesic flow. The propagation of singularities theorem now states in this example that this flow must leave the wave front set $\mathrm{WF}(u)$ invariant, in the sense that $\phi_t^* \mathrm{WF}(u) \subset \mathrm{WF}(u)$. Thus, the propagation of singularities theorem gives information how singularities propagate along the bicharacteristic flow. The theorem as just stated is in fact just a special case of the celebrated Duistermaat-Hörmander propagation of singularities theorem [28], which holds for much more general operators A of real principal type (including e.g. the massive wave equation). The Hamiltonian is then given simply by $h(x,k) = a^{\mu_1 \cdots \mu_N}(x) k_{\mu_1} \dots k_{\mu_N}$ in the general case, where N is the degree of the operator.

Another useful concept in perturbative quantum field theory is that of the scaling degree of a distribution. Let u be a distribution on $\mathbb{R}^n$. The scaling degree, $sd_0(u)$ at the origin of $\mathbb{R}^n$ is defined as

$$sd_0(u) = \inf\{\delta \in \mathbb{R} \mid \lim_{t \to 0+} t^\delta u(tx) = 0\} \tag{501}$$

where the limit is understood in the sense of distributions, i.e., after smearing with a testfunction. One similarly defines the scaling degree $sd_x(u)$ at an arbitrary point x by first translating u by x. On a manifold X, the scaling degree is defined by first localizing u with a cutoff function and then pulling it back with a coordinate chart, $\kappa^*(\psi u)$, as in the definition of the wave-front set. One again verifies that the definition does not depend upon the choice of coordinates.

D Hadamard parametrices

In this appendix, we review the definition of the scalar Hadamard parametrix H^{s}, and the vector Hadamard parametrix, H^{v}, as well as the local expressions for the advanced and retarded propagators in curved spacetime.

D.1 Scalar Hadamard parametrix

In a general curved spacetime, it is not possible to find a closed form expression for $\Delta_{A,R}$, but it is still possible to present a local expression $H_{A,R}$ involving certain recursively defined coefficients, which locally coincides with $\Delta_{A,R}$ modulo C^∞. The distributions $H_{A,R}$ are called "Hadamard parametrices" for $\Delta_{A,R}$. To construct them, let $x, y \in M$, and consider the length functional

$$s(x,y) = \int_a^b \left| g_{\mu\nu}(\gamma(t)) \dot{\gamma}^\mu(t) \dot{\gamma}^\nu(t) \right|^{1/2} dt \tag{502}$$

for C^1-curves $\gamma : [a,b] \to M$ with the property that $\gamma(a) = x$ and $\gamma(b) = y$, which are either spacelike, timelike, or null (but do not switch from one to the other). The functional $s(p,q)$ is invariant under reparametrizations of the curve, so we may choose a parametrization so that $g_{\mu\nu}\dot{\gamma}^\mu\dot{\gamma}^\nu = 1$ along the curve when γ is either spacelike or timelike (such a parameter is called an "affine parameter"). The Euler-Lagrange equations for the functional are then given by

$$\dot{\gamma}^\mu \nabla_\mu \dot{\gamma}^\nu = 0, \tag{503}$$

and curves satisfying this equation are "geodesics". If γ^μ are the components of γ in a local chart, then the geodesic equation reads

$$\ddot{\gamma}^\mu + \Gamma^\mu{}_{\sigma\nu}\dot{\gamma}^\sigma\dot{\gamma}^\nu = 0\,. \tag{504}$$

Two given points x, y may in general be joined by several geodesics, but one can show [61] that every point in M has a neighborhood U such that any pair of points $(x, y) \in U \times U$ may be joined by a unique geodesic lying entirely within U. For $(x, y) \in U \times U$, we define $\sigma(x, y)$ to be the value of the function $\pm s(x, y)^2$ evaluated on the unique geodesic joining x and y, where $+$ is chosen for a spacelike, and $-$ is chosen for a timelike geodesic. In Minkowski spacetime, the function σ is equal to the invariant distance between the points x, y. In any spacetime, the function σ has the important property that

$$g^{\mu\nu}\nabla_\mu\sigma\nabla_\nu\sigma = 4\sigma, \tag{505}$$

where the derivative can act on either the first or second argument. Now let $T : M \to \mathbb{R}$ be a time function. By analogy with flat spacetime, we seek Hadamard parametrices for the advanced and retarded propagators by the following ansatz:

$$H_{A,R}(x,y) = \frac{1}{2\pi}\theta(\mp t(x,y))\Big[u(x,y)\delta(\sigma(x,y)) - v(x,y)\theta(-\sigma(x,y))\Big], \tag{506}$$

Here, u, v are as yet unknown smooth, symmetric functions on $U \times U$ and $t(x, y) = T(x) - T(y)$. This ansatz is consistent with the support properties of the advanced and retarded propagators, and it does not depend on the particular choice of time function. The unknown functions u, v are to be determined imposing in addition the Klein-Gordon equation,

$$\begin{aligned} (\Box - m^2)_x H_{A,R}(x,y) &= \delta(x,y) \quad \text{modulo } C^\infty, & (507)\\ (\Box - m^2)_y H_{A,R}(x,y) &= \delta(x,y) \quad \text{modulo } C^\infty. & (508) \end{aligned}$$

Using the identity (505) one finds that H_A, H_R solve these equations in $U \times U$ modulo C^∞ if the following identities hold for u, v:

$$2\nabla^\mu\sigma\nabla_\mu u = (8 - \Box\sigma)u\,. \tag{509}$$

as well as

$$(\Box - m^2)v = 0, \tag{510}$$

modulo C^∞, and

$$2\nabla^\mu\sigma\nabla_\mu v + (\Box\sigma - 4)v = -(\Box - m^2)u, \quad \text{on } \partial J^\pm(y) \tag{511}$$

where the derivative operators act on the point x. One can show that the unique smooth solution to the equation for u is given by $u = D^{1/2}$, where $D(x, y)$ is the so-called "VanVleck determinant", which is defined as follows. Let $x, y \in U$, and let $A_{\mu\nu} = (\nabla_\mu \otimes \nabla_\nu)\sigma$, so that $A_{\mu\nu}dx^\mu \otimes dy^\nu$

is a tensor in $T^*_x M \otimes T^*_y M$. We can consider the 4-th antisymmetric tensor power of this tensor, which may be viewed as a map

$$\wedge^4 A : \wedge^4 T_x M \to \wedge^4 T^*_y M, \tag{512}$$

where $\wedge^r T_p M$ denotes the space of totally antisymmetric tensors of type $(r,0)$. Clearly, for $r = 4$ this space is 1-dimensional (in 4 dimensions), so if we pick a basis element at points x, y, we can identify $\wedge^4 A$ with a scalar. A choice of the basis element depending only upon the metric (up to a sign) is the Levi-Civita tensor ε. With this choice, D is defined as the scalar obtained from $\wedge^4 A$. In local coordinates,

$$D = \frac{1}{4!} A^{\nu_1}{}_{\mu_1} A^{\nu_2}{}_{\mu_2} A^{\nu_3}{}_{\mu_3} A^{\nu_4}{}_{\mu_4} \varepsilon^{\mu_1\mu_2\mu_3\mu_4} \varepsilon_{\nu_1\nu_2\nu_3\nu_4} \,. \tag{513}$$

where the ε tensors are evaluated at x and y, respectively. While it is not possible to give a similarly explicit solution to the equation for v, it is possible to obtain a solution v in the form of a convergent power series

$$v = \sum_{n=0}^{\infty} v_n \chi(\sigma/\alpha_n) \sigma^n, \tag{514}$$

Here, χ is an arbitrary function of compact support that is equal to 1 in a neighborhood of 0, and $\{\alpha_n\}$ is a sequence growing sufficiently rapidly so as to enforce the convergence of the series. The coefficients are determined recursively as the solutions of the "transport equations"

$$2\nabla_\mu \sigma \nabla^\mu v_0 - (\nabla_\mu \sigma \nabla^\mu \log D - 4) v_0 = -(\Box - m^2) D^{1/2}, \tag{515}$$

from eq. (510) and, for $n > 0$

$$2\nabla_\mu \sigma \nabla^\mu v_n - (\nabla_\mu \sigma \nabla^\mu \log D - 4n - 4) v_n = -\frac{1}{n}(\Box - m^2) v_{n-1} \tag{516}$$

from eq. (511). The solutions to these differential equations are unique if one assumes, as we have done that v_n are smooth (i.e., in particular regular at $x = y$). These solutions can be given in integral form as

$$v_0 = -\frac{1}{2} D^{1/2} \int_0^1 \frac{(\Box - m^2) D^{1/2}}{D^{1/2}} \lambda^2 \, d\lambda \tag{517}$$

and, for $n > 0$

$$v_n = -\frac{1}{2n} D^{1/2} \int_0^1 \frac{(\Box - m^2) v_{n-1}}{D^{1/2}} \lambda^{2n+2} \, d\lambda \tag{518}$$

where the integrand is evaluated at the point $(x(\lambda), y)$, where $x(\lambda) = \mathrm{Exp}_y(\lambda\xi)$, and where $\xi \in T_y M$ is chosen so that $x(1) = x$. Thus, in terms of the Riemannian normal coordinates of x relative to y, then the integrand is thought of as evaluated at the rescaled normal coordinates.

Despite the apparent asymmetry in the construction of u, v, it can be shown that these functions are symmetric in x, y [51, 89], and one shows that, indeed,

$$H_{A,R}(x,y) = \Delta_{A,R}(x,y) \quad \text{modulo } C^\infty \tag{519}$$

in $U \times U$. (It can be proved that exact Greens functions $\Delta_{A,R}$ exist globally, for which the power series expressions therefore define local asymptotic expansions.)

From the advanced and retarded parametrices one can define 2 other parametrices $H_{F,D}$ (for "Feynman" and "Dyson"), given by

$$H_{F,D}(x,y) = \frac{1}{2\pi^2}\Big(\frac{u(x,y)}{\sigma \pm i0} + v(x,y)\log(\sigma \pm i0)\Big) \tag{520}$$

These parametrices are symmetric in x, y. Using the transport equations for u, v, one shows that these, too, are local Green's functions (with δ-function source) modulo C^∞. The wave-front sets of $H_{A,R,F,D}$ are described by the following theorem:

Theorem 7. The wave front set of the 4 Hadamard parametrices are given by

$$\begin{aligned} \mathrm{WF}(H_{A,R}) &= \{(x_1,k_1;x_2,k_2) \mid k_1 \sim -k_2, x_1 \in J^\pm(x_2)\} \\ &\cup \{(x,k;x,-k)\} \end{aligned} \tag{521}$$

$$\begin{aligned} \mathrm{WF}(H_{F,D}) &= \{(x_1,k_1;x_2,k_2) \mid k_1 \sim -k_2, k_1 \in V_\pm^* \,\mathrm{iff}\, x_1 \in J^\pm(x_2)\} \\ &\cup \{(x,k;x,-k)\} \end{aligned} \tag{522}$$

The proof of this theorem is similar to that of the next lemma. It can also be proved that the four parametrices $H_{A,R,F,D}$ are uniquely characterized by their wave front properties. In fact, there is a similar classification of parametrices for any operator or real principal type, as shown by a profound theorem by Duistermaat and Hörmander [28].

In the body of the paper, we use a combination, H, of the above Hadamard parametrices, which is called simply the "local (scalar) Hadamard parametrix" for the operator $\Box - m^2$. It is the distribution on $U \times U$ defined by eq. (154) in terms of the same coefficients u, v that appear above in the local expressions for the advanced and retarded propagators. From identities like

$$\frac{1}{i\pi}\Im\Big(\frac{1}{\sigma + i0t}\Big) = \varepsilon(t)\delta(\sigma), \quad \frac{1}{i\pi}\Im\Big(\log(\sigma + i0t)\Big) = \varepsilon(t)\theta(-\sigma), \tag{523}$$

we get the relations

$$H_F - H_R = -iH = H_A - H_D\,. \tag{524}$$

In view of the symmetry of $H_{F,D}$, there follows the commutator property (532). Furthermore since $H_{A,R,F,D}$ are local Green's functions modulo C^∞ with a δ-function source, there follow the equations of motion

$$(\Box - m^2)_x H(x,y) = 0 \quad \text{modulo } C^\infty, \quad (\Box - m^2)_y H(x,y) = 0 \quad \text{modulo } C^\infty, \tag{525}$$

The local Hadamard parametrix H is important because it characterizes the short distance behavior of any Hadamard state, see Appendix E.

D.2 Vector Hadamard parametrix

The vector Hadamard parametrix $H^{\mathrm{v}}(x,y) = H^{\mathrm{v}}_{\mu\nu}(x,y) dx^\mu \wedge dy^\nu$ is constructed by analogy to the scalar case. It now satisfies the equations

$$(d\delta + \delta d)_x H^{\mathrm{v}}(x,y) = 0 \quad \text{modulo } C^\infty, \quad (d\delta + \delta d)_y H^{\mathrm{v}}(x,y) = 0 \quad \text{modulo } C^\infty, \tag{526}$$

where $\delta = *d*$. In component form, the equations of motion are given by the operator (243). The local vector Hadamard parametrix has an expansion similar to that of the scalar Hadamard parametrix:

$$H^{\mathrm{v}}_{\mu\nu}(x,y) = \frac{1}{2\pi^2}\left(\frac{u_{\mu\nu}(x,y)}{\sigma + i0t} + v_{\mu\nu}(x,y)\log(\sigma + i0t)\right). \tag{527}$$

The coefficients $u_{\mu\nu}, v_{\mu\nu}$ have expansions that are analogous to the scalar case. The quantity $u_{\mu\nu}$ is given explicitly by

$$u_{\mu\nu} = D^{1/2} I_{\mu\nu} \tag{528}$$

where $I : T_x M \to T^*_y M$ is the holonomy of the Levi-civita connection along the unique geodesic connecting x,y ("bitensor of parallel transport"). The expansion coefficients of $v_{\mu\nu}$ as in eq. (514) are again determined by transport equations. The solutions to these equations take exactly the same form as in the scalar case, eq. (518), with the only difference that the scalar Klein-Gordon operator $\Box - m^2$ in those expressions is replaced by the vector wave-operator $g_{\mu\nu}\Box + R_{\mu\nu}$.

E Hadamard states

In the body of the paper, Hadamard 2-point functions play a key role. They were introduced in Sec. 3.1 as bidistributions that are solutions to the wave equation in both entries, that satisfy the commutator property, and that have a certain wave front set. Here we show that these conditions allow one to identify the short distance behavior of any Hadamard 2-point function with that of the local parametrix H introduced in the previous subsection.

Lemma 11. Let $\omega(x,y)$ be a 2-point function of Hadamard form, i.e., the wave front set $\mathrm{WF}(\omega)$ is given by (119). Then locally (i.e., where H is defined), $\omega - H$ is smooth, i.e.,

$$\omega(x,y) = \frac{1}{2\pi^2}\left(\frac{u(x,y)}{\sigma + it0} + v(x,y)\,\log(\sigma + it0)\right) + \quad \text{(smooth function in } x,y). \tag{529}$$

Furthermore, any two Hadamard states can at most differ by a globally smooth function in x,y.

Proof: We first show that, where it is defined, H has a wave front set $\mathrm{WF}(H)$ of Hadamard form, i.e., is given by eq. (119). Since v_i are smooth functions on a convex normal neighborhood, it suffices to prove that $\mathrm{WF}([\sigma + i0t]^{-1})$ and $\mathrm{WF}(\log[\sigma + i0t])$ have the desired form. To determine the wave front set of such distributions, we use the above thm. 6. We apply this theorem to the distributions in question as follows. First, we pick a local coordinate system (ψ, U) in a convex

normal neighborhood U. Within U, we pick a tetrad $e_0, \dots, e_3$ which we use to identify each T_xM with $\mathbb{R}^4$ via the map sending $\xi = (\xi^0, \dots, \xi^3)$ in $\mathbb{R}^4$ to the point $e_x(\xi) = \xi^0 e_0|_x + \dots + \xi^3 e_3|_x$ in T_xM. For each given $x \in U$, we can then write a point $y \in U$ uniquely as $y = \exp_x e_x(\xi)$ for some $\xi \in \mathbb{R}^4$. The mapping $(x,y) \in U \times U \mapsto (\psi(x), \xi)$ thus defines a local coordinate chart in $M \times M$, which we call again ψ. Evidently, it then follows that the pull-back of $(\sigma + i0t)^{-1}$ under ψ is given by the distribution

$$\frac{1}{(y+i0e)^2} = \underset{\eta \in V^+, \eta \to 0}{\mathrm{B.V.}} \frac{1}{(\xi + i\eta)^2}, \tag{530}$$

where $e = (1,0,0,0)$, which is of the form to which we can apply our lemma. Using that the dual cone of the open future lightcone V^+ in Minkowski spacetime is the closure of the past lightcone $\bar{V}^-$, it follows

$$\mathrm{WF}([\sigma + i0t]^{-1}) \subset \psi^*[(\mathbb{R}^4 \times 0) \times (\mathbb{R}^4 \times \bar{V}^-)]. \tag{531}$$

From this, the desired wave front set follows. The logarithmic term is treated in exactly the same fashion. Consider now the distribution $d = \omega - H$. The anti-symmetric part of ω is given by $i\Delta$, and the anti-symmetric part of H is given by

$$H(x,y) - H(y,x) = i\varepsilon(t) \{u(x,y)\delta(\sigma) + v(x,y)\theta(\sigma)\}, \tag{532}$$

where $\varepsilon(t) = 1$ for $t > 0$, and $\varepsilon(t) = -1$ for $t \le 0$. It can be shown that the right side of the equation is equal to $i\Delta$ modulo a smooth function. Thus, $d(x,y)$ is symmetric in x,y modulo a smooth remainder. On the other hand, since we know that H has the same wave front set as ω, we know that

$$\begin{aligned} \mathrm{WF}(d) \subset \; & \{(x_1,k_1,x_2,k_2) \in T^*M \times T^*M; \\ & x_1 \text{ and } x_2 \text{ can be joined by null-geodesic } \gamma \qquad (533) \\ & k_1 = \dot\gamma(0) \text{ and } k_2 = -\dot\gamma(1), \text{ and } k_1 \in \bar{V}^+\}. \qquad (534) \end{aligned}$$

which is evidently not a symmetric set. Thus, the only possibility is that, in fact, $\mathrm{WF}(d) = \emptyset$, meaning that $d \in C^\infty$, or equivalently, that $\omega = H$ modulo smooth. This proves the lemma. □

Another proposition about Hadamard 2-point function underlying the "deformation argument construction" of Hadamard states given in subsection 4.2 is the following:

Theorem 8. Let ω be a positive definite distributional bi-solution such that $\mathrm{WF}(\omega)$ has the Hadamard wave front property in an open neighborhood of $\Sigma \times \Sigma$, where Σ is a Cauchy surface. Then $\mathrm{WF}(\omega)$ has the Hadamard form globally on $M \times M$.

The proof of the theorem is a simple application of the propagation of singularities theorem for solutions of the Klein-Gordon equation described in the previous subsection.

A (quasifree) Hadamard state is a 2-point function that is in addition positive definite, $\omega(\bar{f}, f) \ge 0$ for any testfunction. The positivity implies an even stronger "local-to-global theorem" than the one given above [94]:

Theorem 9. Let ω be a bi-solution to the Klein-Gordon equation in both entries, with anti-symmetric part $i\Delta$, and with the property that any point $x \in M$ has a globally hyperbolic neighborhood N such that $\mathrm{WF}(\omega)$ is of Hadamard form in $N \times N$. Then $\mathrm{WF}(\omega)$ has the Hadamard form globally in $M \times M$.

References

[1] A. Abrahams, A. Anderson, Y. Choquet-Bruhat and J. W. . York, "Einstein And Yang-Mills Theories In Hyperbolic Form Without Gauge Fixing," Phys. Rev. Lett. **75** (1995) 3377 [arXiv:gr-qc/9506072].

[2] L. Alvarez-Gaume and E. Witten, "Gravitational Anomalies," Nucl. Phys. B **234**, 269 (1984).

[3] A. Ashtekar and A. Sen, "On The Role Of Space-Time Topology In Quantum Phenomena: Superselection Of Charge And Emergence Of Nontrivial Vacua," J. Math. Phys. **21**, 526 (1980).

[4] G. Barnich, M. Henneaux, T. Hurth, and K. Skenderis: "Cohomological analysis of gauge-fixed theories," hep-th/9910201

[5] G. Barnich, F. Brandt and M. Henneaux, "Local BRST cohomology in gauge theories," Phys. Rept. **338**, 439 (2000)

[6] I. A. Batalin, G. A. Vilkovisky, Phys. Lett. **B102**, 27 (1981), Nucl. Phys. **B234**, 106 (1984), J. Math. Phys. **26**, 172 (1985)

[7] F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz and D. Sternheimer, "Deformation Theory And Quantization. 2. Physical Applications," Annals Phys. **111**, 111 (1978).

[8] F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz and D. Sternheimer, "Deformation Theory And Quantization. 1. Deformations Of Symplectic Structures," Annals Phys. **111**, 61 (1978).

[9] C. Becchi, A. Rouet and R. Stora, "Renormalization Of The Abelian Higgs-Kibble Model," Commun. Math. Phys. **42**, 127 (1975).

[10] C. Becchi, A. Rouet and R. Stora, "Renormalization Of Gauge Theories," Annals Phys. **98**, 287 (1976).

[11] N. D. Birrell and P. C. W. Davies, "Quantum Fields In Curved Space," *Cambridge, Uk: Univ. Pr. (1982) 340p*

[12] N.N. Bogoliubov and D.V. Shirkov: "Introduction to the Theory of Quantized Fields." John Willey & Sons, Inc., New York, third edition, 1980.

[13] R. Bott and L. Tu, "Differential Forms in Algebraic Topology." Springer-Verlag, 1981.

[14] F. Brennecke: "Zum Anomalie-Problem der Master-Ward-Identität," Diploma Thesis (in German), Hamburg 2005, available at `http://www.desy.de/uni-th/lqp/psfiles/dipl-brennecke.ps.gz`

[15] F. Brennecke and M. Duetsch, "Removal of violations of the Master Ward Identity in perturbative QFT," [hep-th/0705.3160]

[16] [4] Bros, J., Iagolnitzer, D., "Causality and local analyticity:a mathematical study", Ann. Inst H. Poincaré **A 18**, 174 (1973), see also D. Iagolnitzer, Lett. Math. Phys. **21** 323 (1991), and references therein.

[17] R. Brunetti, K. Fredenhagen and M. Kohler, "The microlocal spectrum condition and Wick polynomials of free fields on curved spacetimes," Commun. Math. Phys. **180**, 633 (1996)

[18] R. Brunetti and K. Fredenhagen, "Microlocal analysis and interacting quantum field theories: Renormalization on physical backgrounds," Commun. Math. Phys. **208**, 623 (2000)

[19] R. Brunetti, K. Fredenhagen and R. Verch, "The generally covariant locality principle: A new paradigm for local quantum physics," Commun. Math. Phys. **237**, 31 (2003) [arXiv:math-ph/0112041].

[20] T. S. Bunch, "Bphz Renormalization Of Lambda Phi**4 Field Theory In Curved Space-Time," Annals Phys. **131**, 118 (1981).

[21] A. Connes and D. Kreimer: "Renormalization in quantum field theory and Riemann-Hilbert problem I: The Hopf algebra structure of graphs and the main theorem," Comm. Math. Phys. **210**, 249 (2000)

[22] A. Connes and D. Kreimer: "Renormalization in quantum field theory and Riemann-Hilbert problem II: The beta function, diffeomorphisms, and the renormalization group," Comm. Math. Phys. **216**, 215 (2001)

[23] E.E. Boos and A. I. Davydychev: "A method for calculating massive Feynman diagrams," Theor. Math. Phys. **89**, 1052 (1991) [Teor. Mat. Fiz. **89**, 56 (1991)].

[24] B. S. DeWitt and R. W. Brehme, "Radiation damping in a gravitational field," Annals Phys. **9**, 220 (1960).

[25] B. S. DeWitt: "Dynamical Theory of Groups and Fields" (Les Houches Lectures 1963) (New York: Gordon and Breach)

[26] B. DeWitt: “The Global Approach to Quantum Field Theory,” Oxford University Press, Oxford 2003

[27] B. DeWitt and C. DeWitt-Morette, “From the Peierls bracket to the Feynman functional integral,” Annals Phys. **314**, 448 (2004).

[28] J.J. Duistermaat and L. Hormander, “Fourier integral operators II,” Acta Math., **128** (1972), 183-269.

[29] M. Dütsch, T. Hurth, F. Krahe and G. Scharf, “Causal Construction Of Yang-Mills Theories. 1,” Nuovo Cim. A **106** (1993) 1029.

[30] M. Dütsch, T. Hurth, F. Krahe and G. Scharf, “Causal construction of Yang-Mills theories. 2,” Nuovo Cim. A **107** (1994) 375.

[31] M. Dütsch, T. Hurth and G. Scharf, “Causal Construction Of Yang-Mills Theories. 3,” Nuovo Cim. A **108**, 679 (1995).

[32] M. Dütsch, T. Hurth and G. Scharf, “Causal construction of Yang-Mills theories. 4. Unitarity,” Nuovo Cim. A **108**, 737 (1995).

[33] M. Dütsch, “On gauge invariance of Yang-Mills theories with matter fields,” Nuovo Cim. A **109**, 1145 (1996).

[34] M. Dütsch and K. Fredenhagen, “Causal perturbation theory in terms of retarded products, and a proof of the action Ward identity,” arXiv:hep-th/0403213, to appear in Rev. Math. Phys.

[35] M. Dütsch and K. Fredenhagen, “The master Ward identity and generalized Schwinger-Dyson equation in classical field theory,” Commun. Math. Phys. **243**, 275 (2003) [arXiv:hep-th/0211242].

[36] M. Dütsch and F. M. Boas, “The master Ward identity,” Rev. Math. Phys. **14**, 977 (2002) [arXiv:hep-th/0111101].

[37] M. Dütsch and K. Fredenhagen, “Perturbative algebraic field theory, and deformation quantization,” Published in *Siena 2000, Mathematical physics in mathematics and physics* 151-160, and Fields Inst. Commun. **30**, 151 (2001) arXiv:hep-th/0101079.

[38] M. Dütsch and K. Fredenhagen, “Algebraic quantum field theory, perturbation theory, and the loop expansion,” Commun. Math. Phys. **219**, 5 (2001) [arXiv:hep-th/0001129].

[39] M. Dütsch and K. Fredenhagen, “A local (perturbative) construction of observables in gauge theories: The example of QED,” Commun. Math. Phys. **203**, 71 (1999)

[40] M. Dütsch, "Proof of perturbative gauge invariance for tree diagrams to all orders," Annalen Phys. **14**, 438 (2005) [arXiv:hep-th/0502071].

[41] M. Dubois-Violette, M. Talon and C. M. Viallet, "Brs Algebras: Analysis Of The Consistency Equations In Gauge Theory," Commun. Math. Phys. **102**, 105 (1985).

[42] M. Dubois-Violette, M. Henneaux, M. Talon and C. M. Viallet, "Some results on local cohomologies in field theory," Phys. Lett. B **267**, 81 (1991).

[43] M. Dubois-Violette, M. Henneaux, M. Talon, C.M. Viallet, Phys. Lett. **B289**, 361 (1992)

[44] H. Epstein and V. Glaser: "The rôle of locality in perturbation theory," Ann. Inst. H. Poincaré Sec. **A XIX**, 211–295 (1973)

[45] L. D. Faddeev and V. N. Popov, "Feynman diagrams for the Yang-Mills field," Phys. Lett. B **25** (1967) 29.

[46] C. J. Fewster and M. J. Pfenning, "A quantum weak energy inequality for spin-one fields in curved spacetime," J. Math. Phys. **44**, 4480 (2003)

[47] K. Fredenhagen: "The algebraic theory of superselection sectors," University of Hamburg lecture notes, available at `www.desy.de/uni-th/lqp/psfiles/superselect.ps.gz`

[48] K. Fredenhagen and M. Küskü: Unpublished notes.

[49] S. A. Fulling, "Aspects of quantum field theory in curved spacetime," London Math. Soc. Student Texts **17**, 1 (1989).

[50] S. A. Fulling, F. J. Narcowich and R. M. Wald, "Singularity Structure Of The Two Point Function In Quantum Field Theory In Curved Space-Time. Ii," Annals Phys. **136**, 243 (1981), see also S. A. Fulling, M. Sweeny and R. M. Wald, "Singularity Structure Of The Two Point Function In Quantum Field Theory In Curved Space-Time," Commun. Math. Phys. **63**, 257 (1978).

[51] F.G. Friedlaender: "The Wave Equation on Curved Space-Time," Cambridge University Press, Cambridge 1975.

[52] G. J. Galloway, K. Schleich, D. M. Witt and E. Woolgar, "Topological censorship and higher genus black holes," Phys. Rev. D **60**, 104039 (1999) [arXiv:gr-qc/9902061].

[53] I. M. Gelfand and G. E. Shilov, *Les distributions I*, Dunod, Paris 1972

[54] C. Q. Geng and R. E. Marshak, "Uniqueness of quark and lepton representations in the standard model from the anomalies viewpoint," Phys. Rev. D **39**, 693 (1989).

[55] R. Geroch, "Partial Differential Equations of Physics," arXiv:gr-qc/9602055.

[56] D. R. Grigore: “Ward identities and renormalization of general gauge theories,” J. Phys. A: Math. Gen. **37**, 2803 (2004)

[57] D. R. Grigore: “The structure of the anomalies of gauge theories in the causal approach,” J. Phys. A: Math. Gen. **35**, 1665 (2002)

[58] D. R. Grigore: “On the uniqueness of the non-abelian gauge-theories in Epstein-Glaser approach to renormalization,” Rom. J. Phys. **44**, (1999)

[59] A. Guichardet: “Cohomologie des groupes topologiques et des algebres de Lie,” Textes Mathematiques 2, Cedic/Fernard Nathan, Paris (1980)

[60] M. Henneaux and C. Teitelboim, “Quantization of gauge systems,” Princeton University Press (1992)

[61] S. W. Hawking and G. F. R. Ellis, “The Large scale structure of space-time,” *Cambridge University Press, Cambridge, 1973*

[62] S. Hollands and R. M. Wald, “Local Wick polynomials and time ordered products of quantum fields in curved spacetime,” Commun. Math. Phys. **223**, 289 (2001)

[63] S. Hollands and R. M. Wald, “Existence of local covariant time ordered products of quantum fields in curved spacetime,” Commun. Math. Phys. **231**, 309 (2002)

[64] S. Hollands and R. M. Wald, “On the renormalization group in curved spacetime,” Commun. Math. Phys. **237**, 123 (2003)

[65] S. Hollands and W. Ruan, “The state space of perturbative quantum field theory in curved space-times,” Annales Henri Poincare **3**, 635 (2002) [arXiv:gr-qc/0108032].

[66] S. Hollands and R. M. Wald, “Conservation of the stress tensor in interacting quantum field theory in curved spacetimes,” Rev. Math. Phys. **17**, 227 (2005) [arXiv:gr-qc/0404074].

[67] S. Hollands, “A general PCT theorem for the operator product expansion in curved spacetime,” Commun. Math. Phys. **244**, 209 (2004) [arXiv:gr-qc/0212028].

[68] S. Hollands, “The operator product expansion for perturbative quantum field theory in curved spacetime,” Commun. Math. Phys., in print, arXiv:gr-qc/0605072.

[69] G. ’t Hooft and M. Veltman, “Diagrammar,” CERN-73-09;

[70] G. ’t Hooft and M. J. G. Veltman, “Regularization And Renormalization Of Gauge Fields,” Nucl. Phys. B **44**, 189 (1972).

[71] G. ’t Hooft and M. J. G. Veltman, “Combinatorics of gauge fields,” Nucl. Phys. B **50**, 318 (1972).

[72] L. Hormander, “The analysis of linear partial differential operators, I,” Springer Verlag (1983)

[73] T. Hurth: “Non-abelian gauge symmetry in the causal Epstein-Glaser approach,” Int. J. Mod. Phys. **A12** 4461 (1995)

[74] T. Hurth: “Non-Abelian gauge theories, the causal approach,” Ann. Phys. **244** 340 (1995)

[75] T. Hurth and K. Skenderis: “Quantum Noether method,” Nucl. Phys. **B** 566 (1999)

[76] V. Iyer and R. M. Wald, “Some properties of Noether charge and a proposal for dynamical black hole entropy,” Phys. Rev. D **50**, 846 (1994)

[77] W. Junker: “Hadamard states, adiabatic vacua, and the construction of physical states for scalar quantum fields in curved space-time,” Rev. Math. Phys. **8** (1996) 1091 [Erratum-ibid. **14** (2002) 511]

[78] W. Junker and E. Schrohe: “Adiabatic vacuum states on general spacetime manifolds: Definition, construction, and physical properties,” Annales Poincare Phys. Theor. **3**, 1113 (2002)

[79] G. Källen: “Formal integration of the equations of quantum theory in the Heisenberg representation,” Ark. Fysik **2** 371 (1950)

[80] B. S. Kay and R. M. Wald, “Theorems on the Uniqueness and Thermal Properties of Stationary, Nonsingular, Quasifree States on Space-Times with a Bifurcate Killing Horizon,” Phys. Rept. **207**, 49 (1991).

[81] M. Kohler: “The Stress energy tensor of a locally supersymmetric quantum field on a curved space-time,” (PhD thesis, Hamburg 1995) arXiv:gr-qc/9505014.

[82] C. Kopper and V. F. Muller: “Renormalization proof for massive $phi(4)^4$ theory on Riemannian manifolds,” arXiv:math-ph/0609089.

[83] D. Kreimer: “Anatomy of a gauge theory,” Annals Phys. **321**, 2757 (2006) [arXiv:hep-th/0509135].

[84] T. Kugo and I. Ojima: “Local Covariant Operator Formalism Of Nonabelian Gauge Theories And Quark Confinement Problem,” Prog. Theor. Phys. Suppl. **66**, 1 (1979). T. Kugo and I. Ojima, “Manifestly Covariant Canonical Formulation Of Yang-Mills Field Theories: Physical State Subsidiary Conditions And Physical S Matrix Unitarity,” Phys. Lett. B **73**, 459 (1978).

[85] M. Kusku, “The free Maxwell field in curved spacetime,” Diploma thesis (2001), DESY-THESIS-2001-040

[86] S. Lazzarini and J. M. Garcia-Bondia: “Improved Epstein-Glaser renormalization. II. Lorentz invariant framework,” J. Math. Phys. **44** (2003) 3863-3875

[87] D. M. Marolf: “The Generalized Peierls bracket,” Annals Phys. **236**, 392 (1994) [arXiv:hep-th/9308150].

[88] J. A. Minahan, P. Ramond and R. C. Warner: “A comment on the anomaly cancellation in the Standard Model,” Phys. Rev. D **41**, 715 (1990).

[89] V. Moretti, “Proof of the symmetry of the off-diagonal Hadamard/Seeley-deWitt’s coefficients in C(infinity) Lorentzian manifolds by a ’local Wick rotation’,” Commun. Math. Phys. **212**, 165 (2000) [arXiv:gr-qc/9908068].

[90] O. Piguet and S. Sorella, “Algebraic Renormalization,” Springer Lecture Notes in Physics, (1995)

[91] R. E. Peierls, “The Commutation laws of relativistic field theory,” Proc. Roy. Soc. Lond. A **214**, 143 (1952).

[92] D. Prange: “Lorentz covariance in Epstein-Glaser renormalization,” [arXiv:hep-th/9904136]

[93] M. J. Radzikowski, “Micro-Local Approach To The Hadamard Condition In Quantum Field Theory On Curved Space-Time,” Commun. Math. Phys. **179**, 529 (1996).

[94] M. J. Radzikowski, “A Local to global singularity theorem for quantum field theory on curved space-time,” Commun. Math. Phys. **180**, 1 (1996).

[95] M. Sato, T. Kawai, and M. Kashiwara, "Hyperfunctions and pseudo-differential equations," Lecture Notes in Mathematics 287, 265-529, Springer Verlag (1973); M. Sato: "Theory of hyperfunctions I, II," J. Fac. Sci. Univ. Tokyo t 8, 139-193, 387-437 (1959-1960) MR 22:4951, MR24A:2237

[96] G. Scharf: “Quantum gauge theories: A true ghost story,” Wiley, New York (2001)

[97] G. Scharf: “Finite Quantum Electrodynamics,” Springer, Berlin Heidelberg New York, 1989 and 1995

[98] E. R. Speer, “Analytic Renormalization,” J. Math. Phys. **9**, 1404 (1968)

[99] E. R. Speer, “On the structure of analytic renormalization,” Commun. Math. Phys. **31**, 23 (1971)

[100] E. R. Speer, “Analytic renormalization using many space-time dimensions,” Commun. Math. Phys. **37**, 83 (1974)

[101] O. Steinmann: “Perturbative QED and Axiomatic Field Theory,” Springer Verlag (2000)

[102] R. Stora: “Pedagogical experiments in renormalized perturbation theory,” in conference Theory of Renormalization and Regularization, Hesselberg, Germany (2002), at `http://wwwthep.physik.uni-mainz.de∼scheck/Hessbg02.html`

[103] R. Stora: “Continuum gauge theories,” 1976 Cargese Lectures published in *New developements in quantum field theory and statistical physics,* eds. M. Levy and P. Mitter NATO ASI Series B26 (Plenum, 1977)

[104] R. Stora: “Algebraic structure and Topological Origin of Anomalies,” Seminar at Cargese Summer Inst. Sept. 1-15 1983, published in *Progress in gauge field theory*, eds. ’t Hooft et al. (Plenum, 1984)

[105] W. D. van Suijlekom: “Renormalization of gauge fields: A Hopf algebra approach,” [arXiv:hep-th/0610137]

[106] W. D. van Suijlekom: “The Hopf algebra of Feynman graphs in Quantum Electrodynamics,” Lett. Math. Phys. **77**, 265 (2006)

[107] R. Verch, “A spin-statistics theorem for quantum fields on curved spacetime manifolds in a generally covariant framework,” Commun. Math. Phys. **223**, 261 (2001) [arXiv:math-ph/0102035].

[108] R. M. Wald, “Quantum field theory in curved space-time and black hole thermodynamics,” *Chicago, USA: Univ. Pr. (1994) 205 p*

[109] R. M. Wald, “On identically closed forms locally constructed from a field” J. Math. Phys. **31** (1990) 2378-2384

[110] S. Weinberg: “The Quantum Theory of Fields, Volume II,” Cambridge University Press (1996)

[111] E. Wigner, Ann. Math. **40** 149 (1939)

[112] J. Zinn-Justin, in “Trends in Elementary Particle Theory — International Summer Institute in Theoretical Physics in Bonn 1974” (Springer-Verlag, Berlin, 1975)